

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2003, 17:31:56 ; Search time 2436 Seconds
(without alignments)
11003.154 Million cell updates/sec

Title: US-09-308-397-1

Perfect score: 921

Sequence: 1 atgactaaaacagccttttt.....tagcacttttagaaaaatag 921

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.cm.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.cm.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inrv.*

32: em.htg.other.*

33: em.htg.mu.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.man.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mu.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	921	100.0	111135	2	SPNEU1906	AL449928 Streptococcus
2	913	99.1	921	6	AX194049	AX194049 Sequence
3	913	99.1	11523	1	AF197933	AF197933 Streptococcus
4	913	99.1	12039	1	AE007354	AE007354 Streptococcus
5	913	99.1	19702	6	BD003687	BD003687 Polynucleotide
6	911.4	99.0	10925	1	AE008418	AE008418 Streptococcus
7	908.4	98.6	945	6	AX194275	AX194275 Sequence
8	523.2	56.8	1196	6	AR193756	AR193756 Sequence
9	439.4	47.7	10909	1	AE006603	AE006603 Streptococcus
10	439.4	47.7	50416	1	AE014165	AE014165 Streptococcus
11	437.8	47.5	14456	1	AE010088	AE010088 Streptococcus
12	388.8	42.2	10701	1	AE006311	AE006311 Lactococcus
13	219.6	23.8	14530	1	AE004276	AE004276 Vibrio cholerae
14	212.4	23.1	11845	1	AE013105	AE013105 Thermoanaerobacter
15	204	22.1	950	6	AX432728	AX432728 Sequence
16	202	22.1	960	6	AX432671	AX432671 Sequence
17	200.2	21.7	298050	1	AP003189	AP003189 Clostridium
18	196.4	21.3	10857	1	U32701	U32701 Haemophilus
19	187.2	20.3	12890	1	AE005319	AE005319 Escherichia
20	187.2	20.3	222605	1	AF002555	AF002555 Escherichia
21	185.6	20.2	1440	1	ECFABHDG	Z11565 E.coli fabH
22	185.6	20.2	1608	1	ECOFABD	M87040 Escherichia
23	185.6	20.2	13051	1	AE000210	AE000210 Escherichia
24	185.6	20.2	16379	1	D90745	D90745 Escherichia
25	179.6	19.5	27779	1	BSV13937	Y13937 Bacillus subtilis
26	179.6	19.5	208780	1	BSUB0009	Z99112 Bacillus subtilis
27	179.4	19.5	10684	1	AE006228	AE006228 Pasteurella
28	178.4	19.4	13165	1	AE007854	AE007854 Clostridium
29	170	18.5	6394	1	AB021978	AB021978 Vibrio cholerae
30	168	18.2	4430	1	AF044668	AF044668 Salmonella
31	168	18.2	20097	1	AE008752	AE008752 Salmonella
32	168.8	18.1	303249	1	AP001515	AP001515 Bacillus subtilis
33	165.6	18.0	2301	6	AX415528	AX415528 Sequence
34	165	17.9	13099	1	AE013778	AE013778 Yersinia
35	165	17.9	193050	1	AJ414149	AJ414149 Yersinia
36	164.8	17.9	254050	1	AL627269	AL627269 Salmonella
37	163.4	17.7	110000	2	AC020884	Continuation (2 of AC020884)
38	162.4	17.6	313450	1	AL596170	AL596170 Listeria
39	162.4	17.6	319630	6	AX413016	AX413016 Sequence
40	162.4	17.6	34980	6	AX417046	AX417046 Sequence
41	160.4	17.4	347050	1	AL591981	AL591981 Listeria
42	160.2	17.4	2678	1	BSU59433	BSU59433 Bacillus subtilis
43	147.4	16.0	6941	1	AE010528	AE010528 Fusobacterium
44	139.6	15.2	592	6	AX437137	AX437137 Sequence
45	136.6	14.8	14534	1	AE002541	AE002541 Neisseria

ALIGNMENTS

RESULT 1	SPNEU1906	111135 bp	DNA	linear	HTG 11-JUL-2001
LOCUS	Streptococcus pneumoniae clone G54	*** SEQUENCING IN PROGRESS ***			
DEFINITION	in ordered pieces.				
ACCESSION	AL449928				
VERSION	AL449928.1	GI:11545153			
KEYWORDS	HTG; HTGS PHASE2.				
SOURCE	Streptococcus pneumoniae				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
REFERENCE	1 (bases 1 to 111135)				
AUTHORS	Dopazo,J., Mendoza,A., Herrero,G.J., Caldara,F., Humbert,Y., Friedl,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de				

Francisco, M., Polissi, A., Buell, G., Feger, G., Garcia, E., Peitsch, M. and Garcia-Bustos, J.F.
Annotated draft genomic sequence from a *Streptococcus pneumoniae* type 19F clinical isolate

Microb. Drug Resist. 7 (2), 99-125 (2001)

2135329

11442348

2 (bases 1 to 111135)

Dopazo, J., Mendoza, A., Herrero, J., Caldara, F., Polissi, A., Humbert, Y., Friedli, L., Guerrier, M., Grand-Schenk, E., Gandin, C., de Francisco, M., Buell, G., Feger, G., Garcia, E., Peitsch, M. and Garcia-Bustos, J.F.

Direct Submission

Submitted (31-OCT-2000) Research Department, Glaxo Wellcome, S.A., Severo Ochoa 2, 28760 Tres Cantos, SPAIN

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1. .111135

/organism="Streptococcus pneumoniae"

/serotype="19F"

/db_xref="taxon:1313"

/clone="G54"

BASE COUNT 32875 a 20454 c 25072 g 32680 t 54 others

ORIGIN

Query Match 100.0%; Score 921; DB 2; Length 111135;

Best Local Similarity 100.0%; Pred. No. 1.2e-241;

Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACTTAAACAGCCCTTTTATTCTGTCAGAGTGCCAGTACTAGGATCGGACGG 60

DB 26478 ATGACTTAAACAGCCCTTTTATTCTGTCAGAGTGCCAGTACTAGGATCGGACGG 26537

QY 61 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 120

DB 26538 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 26597

QY 121 GGTATGATTACGTTACTCTCATGATACGGAGAGCAACTCAATCAGCCGGTAT 180

DB 26598 GGTATGATTACGTTACTCTCATGATACGGAGAGCAACTCAATCAGCCGGTAT 26657

QY 181 ACGCAACCAAGCCATCTAGCAGTTCGGTGTCTATCTACCGTTTATTCAGAGAAAGGC 240

DB 26658 ACGCAACCAAGCCATCTAGCAGTTCGGTGTCTATCTACCGTTTATTCAGAGAAAGGC 26717

QY 241 TATCAGCCTGATATGGTGTCTGTCTCTTGAGAAATATCTGCTTGGTGGCAAGC 300

DB 26718 TATCAGCCTGATATGGTGTCTGTCTCTTGAGAAATATCTGCTTGGTGGCAAGC 26777

QY 301 GCGCGCTTGGATTTGAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGAA 360

DB 26778 GCGCGCTTGGATTTGAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGAA 26837

QY 361 GAAGCGCTTCTGTGATCTGCGAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGTC 420

DB 26838 GAAGCGCTTCTGTGATCTGCGAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGTC 26897

QY 421 ATTGAAGAGCGCTGTCAAAAGCTTCTGAACTTGGAGTGGTTACTCCAGGCCAATATAAC 480

DB 26898 ATTGAAGAGCGCTGTCAAAAGCTTCTGAACTTGGAGTGGTTACTCCAGGCCAATATAAC 26957

QY 481 ACACCTGCACAAATCGTATCTCTGGAGAGTGGTTGAGTTGATCGAGCGGTTGAATC 540

DB 26958 ACACCTGCACAAATCGTATCTCTGGAGAGTGGTTGAGTTGATCGAGCGGTTGAATC 27017

QY 541 TTGCAAGAGAGCGGTGCCAAGCTTGTATCTCTTAAGGTGTGAGTCCCTTTTCACACC 600

DB 27018 TTGCAAGAGAGCGGTGCCAAGCTTGTATCTCTTAAGGTGTGAGTCCCTTTTCACACC 27077

QY 601 GCTCTCTTGGAGCGTCTAGCCAGAAACTAGCTGAACACTTAGCTCAGGTAAGTTTTC 660

DB 27078 GCTCTCTTGGAGCGTCTAGCCAGAAACTAGCTGAACACTTAGCTCAGGTAAGTTTTC 27137

QY 661 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 720

DB 27138 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 27197

QY 721 GCTCAGCTCTTGACGGTCTAGGTCAGGCAAGAACCGTTTCGTTTCTATCAAGTATTTGGGTC 780

DB 27198 GCTCAGCTCTTGACGGTCTAGGTCAGGCAAGAACCGTTTCGTTTCTATCAAGTATTTGGGTC 27257

QY 781 ATGCAAGAGCAGCGCATTAAGCAACTTTATCGAGATTGACCGGGGAAAGTCTTTGTCAAGT 840

DB 27258 ATGCAAGAGCAGCGCATTAAGCAACTTTATCGAGATTGACCGGGGAAAGTCTTTGTCAAGT 27317

QY 841 TTGTTTAAAAAATTCATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTA 900

DB 27318 TTGTTTAAAAAATTCATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTA 27377

QY 901 GTAGCAGCTTTTAGAAAAATAG 921

DB 27378 GTAGCAGCTTTTAGAAAAATAG 27398

RESULT 2

AX194049

LOCUS

AX194049

DEFINITION

Sequence 30 from Patent WO0149721.

ACCESSION

AX194049

VERSION

AX194049.1

GI:15211644

KEYWORDS

Streptococcus pneumoniae.

ORGANISM

Streptococcus pneumoniae

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Streptococcus.

REFERENCE

1 (bases 1 to 921)

DOUGHERTY, T.J., Pucci, M.J., Dougherty, B.A., Davison, D.B.,

Brucoleri, R.E. and Thanassi, J.A.

Novel bacterial genes and proteins that are essential for cell

viability and their uses

JOURNAL

Patent: WO 0149721-A 30 12-JUL-2001;

Bristol-Myers Squibb Co. (US)

FEATURES

Location/Qualifiers

1..921

/organism="Streptococcus pneumoniae"

/db_xref="taxon:1313"

BASE COUNT 246 a 193 c 232 g 250 t

ORIGIN

Query Match 99.1%; Score 913; DB 6; Length 921;

Best Local Similarity 99.5%; Pred. No. 1.6e-239;

Matches 916; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACTTAAACAGCCCTTTTATTCTGTCAGAGTGCCAGTACTAGGATCGGACGG 60

DB 1 ATGACTTAAACAGCCCTTTTATTCTGTCAGAGTGCCAGTACTAGGATCGGACGG 60

QY 61 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 120

DB 61 GATTTCATGATCAGTATCCGATTCGTCAGAGTATCGAGCGAGTCAGTGTCTC 120

QY 121 GGTATGATTACGTTACTCTCATGATACGGAGAGCAACTCAATCAGACCCGCTAT 180

DB 121 GGTATGATTACGTTACTCTCATGATACGGAGAGCAACTCAATCAGACCCGCTAT 180

QY 181 ACGCAACCAAGCCATCTAGCAGTTCGGTGTCTATCTACCGTTTATTCAGAGAAAGGC 240

DB 181 ACGCAACCAAGCCATCTAGCAGTTCGGTGTCTATCTACCGTTTATTCAGAGAAAGGC 240

QY 241 TATCAGCCTGATATGGTGTCTGTCTCTTGAGAAATATCTGCTTGGTGGCAAGC 300

DB 241 TATCAGCCTGATATGGTGTCTGTCTCTTGAGAAATATCTGCTTGGTGGCAAGC 300


```
QY 901 GTAGCAGCTTTAGAAAAATAG 921
Db 1957 GTAGCAGCTTTAGAAAAATAG 1977

RESULT 5
LOCUS BD003687 19702 bp DNA linear PAT 31-JAN-2002
DEFINITION Polynucleotide of Streptococcus pneumoniae and sequence.
ACCESSION BD003687
VERSION BD003687.1 GI:18631648
KEYWORDS JP 2001501833-A/7.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 19702)
AUTHORS Kunsch,C.A., Choi,G.H., Dillon,P.J., Rosen,C.A., Bara,S.C.,
Fannon,M. and Dougherty,B.A.
TITLE Polynucleotide of Streptococcus pneumoniae and sequence
JOURNAL Patent: JP 2001501833-A 7 13-FEB-2001;
HUMAN GENOME SCIENCES INC
COMMENT OS Unidentified
PN JP 2001501833-A/7
PD 13-FEB-2001
PF 30-OCT-1997 JP 1998520718
PR 31-OCT-1996 US 60/029960
PI CHARLES A KUNSCH, GIL H CHOI, PATRICK J DILLON, CRAIG A ROSEN, PI
STEVEN C BARASH.
PI MICHAEL FANNON BRIAN A DOUGHERTY
PC C12N15/09,A01K67/027,C07K14/315,C07K16/12,C12N1/15,C12N1/19,
PC C12N1/21,
PC C12N5/10,C12P21/02,C12Q1/68,G06F17/30,C12N15/00,C12N5/00, PC
G06F15/40
CC Strandedness: Double;
CH Topology: Linear;
FH Key Location/Qualifiers
FT source 1..19702
FT location/Qualifiers
FT 1..19702
FT /organism='Unidentified'
FT /organism='unidentified'
FT /db_xref='taxon:32644'

BASE COUNT 5736 a 4548 c 3728 g 5690 t
ORIGIN

Query Match 99.1%; Score 913; DB 6; Length 19702;
Best Local Similarity 99.5%; Pred. NO. 1.8e-239;
Matches 916; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGCTAAACAGCGCTTTTATTTGCTGCTCAAGTCCCGAGTATCTAGGATGGGACGG 60
Db 17149 ATGCTAAACAGCGCTTTTATTTGCTGCTCAAGTCCCGAGTATCTAGGATGGGACGG 17090

QY 61 GATTTCATGATCATGATCCGATTGTCAGAAACGATTGATCGAGCGAGTTCAGGTGCTC 120
Db 17089 GATTTCATGATCATGATCCGATTGTCAGAAACGATTGATCGAGCGAGTTCAGGTGCTC 17030

QY 121 GGTTATGATTACGTTATCTCATGATACGGAAGAACAACTCAATCAGCCCGTAT 180
Db 17029 GGTATGATTTCGGTTATCTCATGATACGGAAGAACAACTCAATCAGCCCGTAT 16970

QY 181 ACGCAACAGGCAATTCAGCGACATTCGGTTGCTATCTACCGTTTATTCGAAGAAAGGCG 240
Db 16969 ACGCAACAGGCAATTCAGCGACATTCGGTTGCTATCTACCGTTTATTCGAAGAAAGGCG 16910

QY 241 TATCAGCCTGATATGTTGCTGTTGCTCTTGGAGAACTCTGCTTGGTGGCAAGC 300
Db 16909 TATCAGCCTGATATGTTGCTGTTGCTCTTGGAGAACTCTGCTTGGTGGCAAGT 16850

QY 301 GCGCCTTGGATTGTTGAAGATGCGGTTGCTTGTAGTAAAGCTGAGCCTATATGGA 360
Db 16849 GCGCCTTGGATTGTTGAAGATGCGGTTGCTTGTAGTAAAGCTGAGCCTATATGGA 16790
```

```
QY 361 GAAGCGGCTCTGCTGACTCTGGCAAGATGGTAGCAGTTCTCAATACGCCAGTACAGGTC 420
Db 16789 GAAGCGGCTCTGCTGACTCTGGCAAGATGGTAGCAGTTCTCAATACGCCAGTACAGGTC 16730

QY 421 ATTGAAGAACCTGTCTCAAAAAGCTTCTGAACCTTGAGTGGTGTACTCCAGCCCACTATAAC 480
Db 16729 ATTGAAGAACCTGTCTCAAAAAGCTTCTGAACCTTGAGTGGTGTACTCCAGCCCACTATAAC 16670

QY 481 ACACCTGCAAAATCGTATCTGCGAGAGTGGTGGAGTGTGATCGAGCGGTGAACCTT 540
Db 16669 ACACCTGCAAAATCGTATCTGCGAGAGTGGTGGAGTGTGATCGAGCGGTGAACCTT 16610

QY 541 TTGCAAGAACGAGGTGCGCAACCGCTTGATTCCTCTTAAGGTGTAGGTCCTTTTACACC 600
Db 16609 TTGCAAGAACGAGGTGCGCAACCGCTTGATTCCTCTTAAGGTGTAGGTCCTTTTACACC 16550

QY 601 GCTCTCTTTGAGCTGTAGCCAGAACTAGCTGAAACTTAGCTCAGTAAAGTATTTTCA 660.
Db 16549 GCTCTCTTTGAGCTGTAGCCAGAACTAGCTGAAACTTAGCTCAGTAAAGTATTTTCA 16490

QY 661 GATTTTACTTGTCCCTAGTGGCAATCAGAAAGCTGTGTGATCAAAAAGAGGACATT 720
Db 16489 GATTTTACTTGTCCCTAGTGGCAATCAGAAAGCTGTGTGATCAAAAAGAGGACATT 16430

QY 721 GCTCAGCTCTTCACGCGTCAGGTCAAGAACCCGTTCTTTCTATGAAAGTATTTGGGTC 780
Db 16429 GCTCAGCTCTTCACGCGTCAGGTCAAGAACCCGTTCTTTCTATGAAAGTATTTGGGTC 16370

QY 781 ATGCAAGAACGAGGATAGCAACTTATCGAGTTGACCGGGAAGTCTTGTCAAGT 840
Db 16369 ATGCAAGAACGAGGATAGCAACTTATCGAGTTGACCGGGAAGTCTTGTCAAGT 16310

QY 841 TTGTTTAAAAAATGATCAAACTGCTCAGTCTAGCTCATGTGGAAGATCAACGAGTTTA 900
Db 16309 TTGTTTAAAAAATGATCAAACTGCTCAGTCTAGCTCATGTGGAAGATCAACGAGTTTA 16250

QY 901 GTAGCAGCTTTAGAAAAATAG 921
Db 16249 GTAGCAGCTTTAGAAAAATAG 16229

RESULT 6
LOCUS AE008418 10925 bp DNA linear BCT 13-SEP-2001
DEFINITION Streptococcus pneumoniae R6 section 34 of 184 of the complete genome.
ACCESSION AE008418 AB007317
VERSION AE008418.1 GI:15457935
KEYWORDS Streptococcus pneumoniae R6.
SOURCE Streptococcus pneumoniae R6
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.
REFERENCE 1 (bases 1 to 10925)
AUTHORS Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmore,R., Glass,J.S., Khoja,H., Kraft,A., Lagace,R., LeBlanc,D.J., Lee,L.N., Leikowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenry,M., McLeaster,K., Mundy,C., Nicas,I.I., Norris,F.H., O'Garra,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostek,P.R. Jr., Skatrud,P.L. and Glasz,J.I.
TITLE Genome of the bacterium Streptococcus pneumoniae strain R6
JOURNAL J. Bacteriol. 193 (19), 5709-5717 (2001)
MEDLINE 21425245
PUBMED 11544234
REFERENCE 2 (bases 1 to 10925)
AUTHORS Hoskins,J.A., Alborn,W. Jr., Arnold,J., Blaszcak,L., Burgett,S., DeHoff,B.S., Estrem,S., Fritz,L., Fu,D.-J., Fuller,W., Geringer,C., Gilmore,R., Glass,J.S., Han,A., Khoja,H., Kraft,A., Lagace,R., LeBlanc,D.J., Lee,L.N., Leikowitz,E.J., Lu,J., Matsushima,P., McAhren,S., McHenry,M., McLeaster,K., Mundy,C., Nicas,I.I., Norris,F.H., O'Garra,M., Peery,R., Robertson,G.T., Rockey,P., Sun,P.-M., Winkler,M.E., Yang,Y., Young-Bellido,M., Zhao,G., Zook,C., Baltz,R.H., Jaskunas,S.R., Rostek,P.R. Jr., Skatrud,P.L. and Glasz,J.I.
```

Norris, F. H., O'Gara, M., Peery, R., Robertson, G. T., Rockey, P., Sun, P.-M., Winkler, M. E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C., Baltz, R. H., Jaskunas, S. R., Rosteck, P. R. Jr., Skatrud, P. L., and Glass, J. I.

TITLE	Direct Submission
1. Title Page	1. Title Page
2. Abstract	2. Abstract
3. Introduction	3. Introduction
4. Methods	4. Methods
5. Results	5. Results
6. Discussion	6. Discussion
7. Conclusion	7. Conclusion
8. References	8. References
9. Appendix	9. Appendix
10. Bibliography	10. Bibliography
11. Glossary	11. Glossary
12. Index	12. Index
13. Table of Contents	13. Table of Contents
14. Figures	14. Figures
15. Tables	15. Tables
16. Summary	16. Summary
17. Acknowledgments	17. Acknowledgments
18. Declaration of Interest	18. Declaration of Interest
19. Funding Source	19. Funding Source
20. Author Biographies	20. Author Biographies
21. Correspondence	21. Correspondence
22. References	22. References
23. Appendix	23. Appendix
24. Bibliography	24. Bibliography
25. Glossary	25. Glossary
26. Index	26. Index
27. Table of Contents	27. Table of Contents
28. Figures	28. Figures
29. Tables	29. Tables
30. Summary	30. Summary
31. Acknowledgments	31. Acknowledgments
32. Declaration of Interest	32. Declaration of Interest
33. Funding Source	33. Funding Source
34. Author Biographies	34. Author Biographies
35. Correspondence	35. Correspondence
36. References	36. References
37. Appendix	37. Appendix
38. Bibliography	38. Bibliography
39. Glossary	39. Glossary
40. Index	40. Index
41. Table of Contents	41. Table of Contents
42. Figures	42. Figures
43. Tables	43. Tables
44. Summary	44. Summary
45. Acknowledgments	45. Acknowledgments
46. Declaration of Interest	46. Declaration of Interest
47. Funding Source	47. Funding Source
48. Author Biographies	48. Author Biographies
49. Correspondence	49. Correspondence
50. References	50. References
51. Appendix	51. Appendix
52. Bibliography	52. Bibliography
53. Glossary	53. Glossary
54. Index	54. Index
55. Table of Contents	55. Table of Contents
56. Figures	56. Figures
57. Tables	57. Tables
58. Summary	58. Summary
59. Acknowledgments	59. Acknowledgments
60. Declaration of Interest	60. Declaration of Interest
61. Funding Source	61. Funding Source
62. Author Biographies	62. Author Biographies
63. Correspondence	63. Correspondence
64. References	64. References
65. Appendix	65. Appendix
66. Bibliography	66. Bibliography
67. Glossary	67. Glossary
68. Index	68. Index
69. Table of Contents	69. Table of Contents
70. Figures	70. Figures
71. Tables	71. Tables
72. Summary	72. Summary
73. Acknowledgments	73. Acknowledgments
74. Declaration of Interest	74. Declaration of Interest
75. Funding Source	75. Funding Source
76. Author Biographies	76. Author Biographies
77. Correspondence	77. Correspondence
78. References	78. References
79. Appendix	79. Appendix
80. Bibliography	80. Bibliography
81. Glossary	81. Glossary
82. Index	82. Index
83. Table of Contents	83. Table of Contents
84. Figures	84. Figures
85. Tables	85. Tables
86. Summary	86. Summary
87. Acknowledgments	87. Acknowledgments
88. Declaration of Interest	88. Declaration of Interest
89. Funding Source	89. Funding Source
90. Author Biographies	90. Author Biographies
91. Correspondence	91. Correspondence
92. References	92. References
93. Appendix	93. Appendix
94. Bibliography	94. Bibliography
95. Glossary	95. Glossary
96. Index	96. Index
97. Table of Contents	97. Table of Contents
98. Figures	98. Figures
99. Tables	99. Tables
100. Summary	100. Summary
101. Acknowledgments	101. Acknowledgments
102. Declaration of Interest	102. Declaration of Interest
103. Funding Source	103. Funding Source
104. Author Biographies	104. Author Biographies
105. Correspondence	105. Correspondence
106. References	106. References
107. Appendix	107. Appendix
108. Bibliography	108. Bibliography
109. Glossary	109. Glossary
110. Index	110. Index
111. Table of Contents	111. Table of Contents
112. Figures	112. Figures
113. Tables	113. Tables
114. Summary	114. Summary
115. Acknowledgments	115. Acknowledgments
116. Declaration of Interest	116. Declaration of Interest
117. Funding Source	117. Funding Source
118. Author Biographies	118. Author Biographies
119. Correspondence	119. Correspondence
120. References	120. References
121. Appendix	121. Appendix
122. Bibliography	122. Bibliography
123. Glossary	123. Glossary
124. Index	124. Index
125. Table of Contents	125. Table of Contents
126. Figures	126. Figures
127. Tables	127. Tables
128. Summary	128. Summary
129. Acknowledgments	129. Acknowledgments
130. Declaration of Interest	130. Declaration of Interest
131. Funding Source	131. Funding Source
132. Author Biographies	132. Author Biographies
133. Correspondence	133. Correspondence
134. References	134. References
135. Appendix	135. Appendix
136. Bibliography	136. Bibliography
137. Glossary	137. Glossary
138. Index	138. Index
139. Table of Contents	139. Table of Contents
140. Figures	1

Submitted (27-JUL-2001) Infectious Diseases Research, Eli Lilly and Company, Lilly Research Labs, Indianapolis, IN 46285-0438, USA

[illegible]

source

1. 10925

/organism="Streptococcus pneumoniae R6"

```
/strain="R6"
```

```
/db xref="taxon:171101"
```


KSINTACSSNDAIGDAPRSIKGPFQDVMVVGTEBASIPTPAIAGQALTAISTEDDP
TRASIPDKRNGRVMEGSGMVLSEAEKRGATILAEVVGIGYNTCDAYHTSPH
PEGGAIKAIKLAESEAEISPEQVAYVYNAHGTSTPANKEGSGAIYAVLUGKEVPSST
KSTGHLGAGAVEATVITTEAMRNFVPMTAGTSEVSDYIEANVYVGGGLEKEPIYA
ISNTFGFGHNAVLAFKRWNR"
6866..7351
/gene="accB"
/note="spr0383"
6866..7351
/gene="accB"
/EC_number="6.4.1.2"
/codon_start=1
/transl_table=11
/label="spr0383"
/product="Biotin carboxyl carrier protein of acetyl-CoA
carboxylase"
/protein_id="AAK99187.1"
/db_xref="GI:15457944"
TOVAPALATPSPVAPTSAPAEVPAEASVATEGNLVEPLVGVVYLAAGP
DKPAFTVGDVSRKQTLVIIEMKVNNEIPAPKDGVTTEILVSNEEMVEFGKGLVRI
K"
7348..7770
/gene="fabZ"
Query Match 99.0%; Score 911.4; DB 1; Length 10925;
Best Local Similarity 99.3%; Pred. No. 4.8e-239;
Matches 915; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 ATGACTAAACAGCGCTTTTATTGCTGGTCAAGGTGCCAGATCTAGGATGGACGG 60
Db 3921 ATGACTAAACAGCGCTTTTATTGCTGGTCAAGGTGCCAGATCTAGGATGGACGG 60
Qy 61 GATTTCATGATCAGTATCCGATTGTCACAAAGAACGATTCATCGAGCGATCAGGTGCTC 120
Db 3981 GATTTCATGATCAGTATCCGATTGTTAAAGAACGATTCATCGAGCGATCAGGTGCTA 4040
Qy 121 GGTATGATTACGTTATCTATCATGATACGAGAGAGACAACTCAATCAGACCCGCTAT 180
Db 4041 GGTATGATTGCGTTATCTATCATGATACGAGAGAGACAACTCAATCAGACCCGCTAT 4100
Qy 181 ACGCAACAGCGCTTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240
Db 4101 ACGCAACAGCGCTTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 4160
Qy 241 TATCAGCGTATATGTTGCTGTTCTCTCTGGAGATATCTCTGCTTGGTGGCAAGC 300
Db 4161 TATCAGCGTATATGTTGCTGTTCTCTCTGGAGATATCTCTGCTTGGTGGCAAGC 4220
Qy 301 GCGCCTTGATTTGAAGATCGGTTGCTTGGTCAAGGTGCCAGTATCTAGGATGGACGG 360
Db 4221 GCGCCTTGATTTGAAGATCGGTTGCTTGGTCAAGGTGCCAGTATCTAGGATGGACGG 4280
Qy 361 GAAGCGCTCTGCTGACTCTGCAAGATGGTATGCTCAATACGCGAGTAGAGGTC 420
Db 4281 GAAGCGCTCTGCTGACTCTGCAAGATGGTATGCTCAATACGCGAGTAGAGGTC 4340
Qy 421 ATTGAAGAAGCGCTGCAAAAGCTTCTGAACCTGGAGTGGTACTCCAGCCACTATAC 480
Db 4341 ATTGAAGAAGCGCTGCAAAAGCTTCTGAACCTGGAGTGGTACTCCAGCCACTATAC 4400
Qy 481 ACACCTGCACAAATCGTCATTGCTGGAGAGTGGTGCAGTTGATCGAGCGGTGAACIT 540
Db 4401 ACACCTGCACAAATCGTCATTGCTGGAGAGTGGTGCAGTTGATCGAGCGGTGAACIT 4460
Qy 541 TTGCAAGAAGCGGTGCCAAAGCTTGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
Db 4461 TTGCAAGAAGCGGTGCCAAAGCTTGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4520
Qy 601 GCT 660
Db 4521 TCT 4580

Qy 661 GATTTTACTTGTCCCTAGTGGCAATACAGAGCTGCTGTGATGCACAAAGAGACATT 720
Db 4581 GATTTTACTTGTCCCTAGTGGCAATACAGAGCTGCTGTGATGCACAAAGAGACATT 4640
Qy 721 GCTCAGCTCTTGTACCGCTCAGGTCAAGGAAACCGCTTCCTTTCTATGAAAGTATTGGGTC 780
Db 4641 GCTCAGCTCTTGTACCGCTCAGGTCAAGGAAACCGCTTCCTTTCTATGAAAGTATTGGGTC 4700
Qy 781 ATGCAAGAAGCGGATTAAGCACTTTATCGAGATTGACCGGGGAAAGTCTTGTCAAGT 840
Db 4701 ATGCAAGAAGCGGATTAAGCACTTTATCGAGATTGACCGGGGAAAGTCTTGTCAAGT 4760
Qy 841 TTTGTTAAAAAATGATCAAACTGCTCAGTCTAGTCTCATGTTGGAAGATCAAGCGAGTTA 900
Db 4761 TTTGTTAAAAAATGATCAAACTGCTCAGTCTAGTCTCATGTTGGAAGATCAAGCGAGTTA 4820
Qy 901 GTAGCACTTTTAGAAAAATAG 921
Db 4821 GTAGCACTTTTAGAAAAATAG 4841
RESULT 7
AX194275 945 bp DNA linear PAT 15-AUG-2001
LOCUS Sequence 256 from Patent WO0149721.
DEFINITION AX194275
ACCESSION AX194275
VERSION AX194275.1 GI:15211757
KEYWORDS Streptococcus pneumoniae.
SOURCE Streptococcus pneumoniae.
ORGANISM Streptococcus pneumoniae.
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 945)
AUTHORS Dougherty,T.J., Pucci,M.J., Dougherty,B.A., Davison,D.B.,
Brucoleri,R.E. and Thanassi,J.A.
TITLE Novel bacterial genes and proteins that are essential for cell
viability and their uses
JOURNAL Patent: WO 0149721-A 256 12-JUL-2001;
Bristol-Myers Squibb Co. (US)
FEATURES
Location/Qualifiers
1..945
/organism="Streptococcus pneumoniae"
/db_xref="taxon:1313"
BASE COUNT 254 a 205 c 232 g 254 t
ORIGIN
Query Match 98.6%; Score 908.4; DB 6; Length 945;
Best Local Similarity 99.3%; Pred. No. 2.9e-238;
Matches 912; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 ATGACTAAACAGCGCTTTTATTGCTGGTCAAGGTGCCAGTATCTAGGATGGACGG 60
Db 1 ATGACTAAACAGCGCTTTTATTGCTGGTCAAGGTGCCAGTATCTAGGATGGACGG 60
Qy 61 GATTTTCTATGATCAGTATCCGATTGTCACAAAGAACGATTCATCGAGCGAGTCAAGTGTCT 120
Db 61 GATTTTCTATGATCAGTATCCGATTGTTAAAGAACGATTCATCGAGCGAGTCAAGTGTCTA 120
Qy 121 GGTATGATTTAGTTATCTCATGATACGGAAGAGCAAACTCAATCAGACCCGCTAT 180
Db 121 GGTATGATTTAGTTATCTCATGATACGGAAGAGCAAACTCAATCAGACCCGCTAT 180
Qy 181 ACGCAACAGCGCTTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240
Db 181 ACGCAACAGCGCTTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240
Qy 241 TATCAGCGTATATGTTGCTGTTTGTCTCTTGGAGAAATCTCTGCTTGGTGGCAAGC 300
Db 241 TATCAGCGTATATGTTGCTGTTTGTCTCTTGGAGAAATCTCTGCTTGGTGGCAAGC 300
Qy 301 GCGCCTTGATTTGAAGATCGGTTGCTTGGTATGCTAAGCGTGAAGCGCTTATATGAA 360
Db 301 GCGCCTTGATTTGAAGATCGGTTGCTTGGTATGCTAAGCGTGAAGCGCTTATATGAA 360

```

QY 361 GAAGGGCTCTGCTGACTCTCGGAAGATGGTAGCAGTTCTCAATACCCAGTAGAGGTC 420
Db 361 GAAGGGCTCTGCTGACTCTCGGAAGATGGTAGCAGTTCTCAATACCCAGTAGAGGTC 420
QY 421 ATTGAAGAGCCTGTCAAAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGGCCAACTATAAC 480
Db 421 ATTGAAGAGCCTGTCAAAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGGCCAACTATAAC 480
QY 481 ACACCTGCACAAATCGTCAATCTCTGGAAGAGTGGTTGAGTGTATCGAGCGGTTGAACCT 540
Db 481 ACACCTGCACAAATCGTCAATCTCTGGAAGAGTGGTTGAGTGTATCGAGCGGTTGAACCT 540
QY 541 TTGCAAGAGCAGGTGCCAAAGCGTTGATTCCTCTTTAAGGTGTCAAGTCCCTTTCACACC 600
Db 541 TTGCAAGAGCAGGTGCCAAAGCGTTGATTCCTCTTTAAGGTGTCAAGTCCCTTTCACACC 600
QY 601 GCTCTCCTTGAGCCTGTAGCAGAACTAGCTGAAACTCTAGCTAGCTAGCTAGCTAGCTAGCT 660
Db 601 GCTCTCCTTGAACTGTAGCAGAACTAGCTGAAACTCTAGCTAGCTAGCTAGCTAGCTAGCT 660
QY 661 GATTTTACTTGTCCCTAGTGGCAATACAGAGCTGTGTGATGCAAAAGAGGACATT 720
Db 661 GATTTTACTTGTCCCTAGTGGCAATACAGAGCTGTGTGATGCAAAAGAGGACATT 720
QY 721 GCTCAGCTCTTGACGCGTCAGCTCAAGAAACCCGTTCTCTTATGAAAGTATTGGGGTC 780
Db 721 GCTCAGCTCTTGACGCGTCAGCTCAAGAAACCCGTTCTCTTATGAAAGTATTGGGGTC 780
QY 781 ATGCAAGAGCAGGATAAGCACTTTATCGAGATTGACCGGGGAAAGTCTTGTCAAGT 840
Db 781 ATGCAAGAGCAGGATAAGCACTTTATCGAGATTGACCGGGGAAAGTCTTGTCAAGT 840
QY 841 TTTGTTAAAAAATTGATCAAACTGCTCATTGCTCATTGCTGGAAGATCAAGCGAGTTTA 900
Db 841 TTTGTTAAAAAATTGATCAAACTGCTCATTGCTCATTGCTGGAAGATCAAGCGAGTTTA 900
QY 901 GTAGCATTTTAGAAAA 918
Db 901 GTAGCATTTTAGAAAA 918

RESULT 8
LOCUS AR193756 1196 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 8 from patent US 6348328.
ACCESSION AR193756
VERSION AR193756.1 GI:20240348
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1196)
AUTHORS Black, M. Terence., Hodgson, J. Edward., Knowles, D. Justin, Charles.,
Nicholas, R. Oakley, and Stodola, R. King.
TITLE Compounds
JOURNAL Patent: US 6348328-A 8 19-FEB-2002;
MEDLINE Location/Qualifiers
PUBMED 1. .1196
REFERENCE 355 a 223 c 286 g 332 t
AUTHORS /organism="unknown"
FEATURES
BASE COUNT 355 a 223 c 286 g 332 t
ORIGIN
Query Match 56.8%; Score 523.2; DB 6; Length 1196;
Best Local Similarity 99.1%; Pred. No. 1.6e-132;
Matches 547; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
QY 372 TGTGACTCTGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAGGTCATTGAAGAGC 431
Db 1 TGTGAACTTGGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAGGTCATTGAAGAGC 60
QY 432 CTGTCAAAAAGCTTCT-GNACTGTGAGTGGTTACTCCAGCCCACTATTAACACCTGCAC 490

```

```

Db 61 CTGTCAAAAAGCTTCTGGAACCTTGGAGTGGTTACTCCAGCCCACTATAACACACCTGCAC 120
QY 491 AAATCTCATTTCTCGAGAAAGTGGTTGCAAGTTGATCGACGGTTGAACCTTTTGCAGAAG 550
Db 121 AAATCTCATTTCTCGAGAAAGTGGTTGCAAGTTGATCGACGGTTGAACCTTTTGCAGAAG 180
QY 551 CAGTCCCAAAAGCTTGAATTCCTCTTAAGTGTCAAGTCCCTTTCACACCGCTCTCCTTG 610
Db 181 CAGTCCCAAAAGCTTGAATTCCTCTTAAGTGTCAAGTCCCTTTCACACCGCTCTCCTTG 240
QY 611 AGCTCTCTAGCCAGAACTAGCTGAAACTCTAGCTCAGCTAAAGTTTTCAGATTTTACTT 570
Db 241 AGCTCTCTAGCCAGAACTAGCTGAAACTCTAGCTCAGCTAAAGTTTTCAGATTTTACTT 300
QY 671 GTCCCTCTAGCCCAATACAGAGCTGTCTGATGCAAAAGAGGACATTGCTCAGCTCT 730
Db 301 GTCCCTCTAGCCCAATACAGAGCTGTCTGATGCAAAAGAGGACATTGCTCAGCTCT 360
QY 731 TGACGGCTCAGGTCAAGAAACCCGTTTCGTTTCTATGAAAGTATTGGGGTCATGCAAGAAG 790
Db 361 TGACGGCTCAGGTCAAGAAACCCGTTTCGTTTCTATGAAAGTATTGGGGTCATGCAAGAAG 420
QY 791 CAGGCATAGCAACTTTA-TGAGATTGACCGGGGAAAGTCTTGTCAAGTTTGTATAA 849
Db 421 CAGGCATAGCAACTTTA-TGAGATTGACCGGGGAAAGTCTTGTCAAGTTTGTATAA 480
QY 850 AAAATTTGATCAAACTGCTCATTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGCATT 909
Db 481 AAAATTTGATCAAACTGCTCATTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGCATT 540
QY 910 TTAGAAAAATAG 921
Db 541 TTAGAAAAATAG 552

RESULT 9
LOCUS AE006603/c 10909 bp DNA linear BCT 01-JUN-2001
DEFINITION Streptococcus pyogenes M1 GAS strain SF370, section 132 of 167 of
the complete genome.
ACCESSION AE006603 AE004092
VERSION AE006603.1 GI:13622797
KEYWORDS Streptococcus.
SOURCE Streptococcus pyogenes M1 GAS.
ORGANISM Streptococcus pyogenes M1 GAS.
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
REFERENCE 1 (bases 1 to 10909)
AUTHORS Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K.,
Primeaux, C., Sezate, S., Surorov, A.N., Kenton, S., Lai, H., Lin, S.,
Qian, Y., Jia, H.G., Nejar, F.Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S.W., Roe, B.A., and McLaughlin, R.E.
Complete genome sequence of an M1 strain of Streptococcus pyogenes
Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)
21192684
PUBMED 11296296
REFERENCE 2 (bases 1 to 10909)
AUTHORS Ferretti, J.J., McShan, W.M., Adjic, D., Savic, D., Savic, G., Lyon, K.,
Primeaux, C., Sezate, S., Surorov, A.N., Kenton, S., Lai, H., Lin, S.,
Qian, Y., Jia, H.G., Nejar, F.Z., Ren, Q., Zhu, H., Song, L., White, J.,
Yuan, X., Clifton, S.W., Roe, B.A., and McLaughlin, R.E.
Direct Submission
Submitted (10-APR-2001) Department of Microbiology and Immunology,
University of Oklahoma Health Sciences Center, 940 SL Young Blvd.,
Oklahoma City, OK 73104 USA
FEATURES
source
1. .10909
/organism="Streptococcus pyogenes M1 GAS"
/strain="SF370"
/serotype="M1"
/db xref="taxon:160490"
complement(142..912)
/gene="acca"
gene

```

```

CDS
/note="SPY1743"
/complement(142..912)
/gene="accA"
/EC_number="6.4.1.2"
/note="Best Blastp hit = dbj|BAB06884.1| (AP001517)
acetyl-CoA carboxylase carboxyltransferase alpha subunit
[Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase alpha subunit"
/protein_id="AAK34487.1"
/db_xref="GI:13622798"
/translation="MTDVSRLKEARDQGRLLTLDVANIIFDDFMELHGRHFSDDGA
IVGLAYACQPVTVIGIQKQLQDLNARFQCPNPEGYRKALRLMKQAEKFGFRV
TFNTAGYVGAERGGQGEATKQIMEMSDLPVIAIIIGEGSGGALALAVADQ
VWLNMTYAVLSPGPFASILWKDGGRAEAAELMKITAGELYKMGIVDRIIPEHGYF
SSEIVDIKANLIEQITSLQAKPLDQLLDERYQRFKY"
/complement(909..1775)
/gene="accD"
/note="SPY1744"
/complement(909..1775)
/gene="accD"
/note="Best Blastp hit = sp|054776|ACCD_SYPN7
ACTYL-COENZYME A CARBOXYLASE CARBOXYL TRANSFERASE SUBUNIT
BETA [ACCase BETA CHAIN] >gi|1399855|gb|AA82034.1|
(U59237) carboxyltransferase beta subunit [Synechococcus
PCC7942]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase beta subunit"
/protein_id="AAK34488.1"
/db_xref="GI:13622799"
/translations="MALFRKDKYIRITPNNSLKGVSNNVPEVDELFAKCPACKHM
IYKDLGLAKICTPCSTNFRISAQERLLTVDGSEFQELTSDTDLPLFPYQFKL
QKAKETGLHEAVLTGKAMVKEQKIALIMDSHFIMASGTVVGEKITLPLFLAELEN
LPVVTFTAGSARMQEIGSLMQAKVSAVKHNSAGLFTILTDTPTGGVTASFA
MEGDIILAEQPSQLVGFAGRRVIETTVRENLPDFOKAELQDHGFVDAIVKRTLRDK
IAHLVAFHGGGQ"
/complement(1784..3148)
/gene="accC"
/note="SPY1745"
/complement(1784..3148)
/gene="accC"
/note="The N-terminal amino acid sequence of this ORF has
been determined from a spot isolated by 2-D gel
electrophoresis from another strain of S. pyogenes. Hogan,
D. A., Whitton, M. M., Rogers, J. and R. A. VanBogelen.
2000. Two-dimensional gel electrophoresis map of
Streptococcus pyogenes proteins. Unpublished data.; Best
Blastp hit = pir|A69581|acetyl-CoA carboxylase (biotin
carboxylase subunit) accC - Bacillus subtilis
>gi|1303913|dbj|BAAL2569.1| (D84432) ybhX [Bacillus
subtilis] >gi|2634868|emb|CAB14365.1| (Z99116) acetyl-CoA
carboxylase subunit (biotin carboxylase subunit) [Bacillus
subtilis]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase biotin
carboxylase subunit"
/protein_id="AAK34489.1"
/db_xref="GI:13622800"
/translations="MFKKILIANRGEIAVRIIRAAARELGISTVAVYSEADKEALHTIL
ADRAICIGARSRESYLNNSVLSAAIVTGAQAIHPGFLSENSKFATMCEEMNKF
IGPSASVMDKMGKINARSEMIAKAGVPVPGSDGVYNAQEAALANKIGVPMUKAS
AGGGGKIRKVEADLEAFNAASCEALGAFNGAMYLEKVIYPARHTEVQILDAY
GNIIHIGERDCSILORNNQVLRPEPSIAGNTLRLHMGQAARAAEAVANAGTIEF
LLDESEKVFEMNTRIQUERHPVFTGVDIVKEQIKAAQGLANQEDITIGH
ALBCRNAENTAFNPFSFGKITDLYMPGSGVGLRVDVSAVNGYAIPIPYDSMIAKII
VHGSNFDALMKRQALVELEIEGITITDQDLDTLSKRKRVAGDYDTSFLMETFLPH
YNKD"
/complement(3180..3599)
/gene="fabZ"
/note="SPY1746"

```

```

CDS
/note="SPY1743"
/complement(142..912)
/gene="accA"
/EC_number="6.4.1.2"
/note="Best Blastp hit = dbj|BAB06884.1| (AP001517)
acetyl-CoA carboxylase carboxyltransferase alpha subunit
[Bacillus halodurans]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase alpha subunit"
/protein_id="AAK34487.1"
/db_xref="GI:13622798"
/translation="MTDVSRLKEARDQGRLLTLDVANIIFDDFMELHGRHFSDDGA
IVGLAYACQPVTVIGIQKQLQDLNARFQCPNPEGYRKALRLMKQAEKFGFRV
TFNTAGYVGAERGGQGEATKQIMEMSDLPVIAIIIGEGSGGALALAVADQ
VWLNMTYAVLSPGPFASILWKDGGRAEAAELMKITAGELYKMGIVDRIIPEHGYF
SSEIVDIKANLIEQITSLQAKPLDQLLDERYQRFKY"
/complement(909..1775)
/gene="accD"
/note="SPY1744"
/complement(909..1775)
/gene="accD"
/note="Best Blastp hit = sp|054776|ACCD_SYPN7
ACTYL-COENZYME A CARBOXYLASE CARBOXYL TRANSFERASE SUBUNIT
BETA [ACCase BETA CHAIN] >gi|1399855|gb|AA82034.1|
(U59237) carboxyltransferase beta subunit [Synechococcus
PCC7942]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase beta subunit"
/protein_id="AAK34488.1"
/db_xref="GI:13622799"
/translations="MALFRKDKYIRITPNNSLKGVSNNVPEVDELFAKCPACKHM
IYKDLGLAKICTPCSTNFRISAQERLLTVDGSEFQELTSDTDLPLFPYQFKL
QKAKETGLHEAVLTGKAMVKEQKIALIMDSHFIMASGTVVGEKITLPLFLAELEN
LPVVTFTAGSARMQEIGSLMQAKVSAVKHNSAGLFTILTDTPTGGVTASFA
MEGDIILAEQPSQLVGFAGRRVIETTVRENLPDFOKAELQDHGFVDAIVKRTLRDK
IAHLVAFHGGGQ"
/complement(1784..3148)
/gene="accC"
/note="SPY1745"
/complement(1784..3148)
/gene="accC"
/note="The N-terminal amino acid sequence of this ORF has
been determined from a spot isolated by 2-D gel
electrophoresis from another strain of S. pyogenes. Hogan,
D. A., Whitton, M. M., Rogers, J. and R. A. VanBogelen.
2000. Two-dimensional gel electrophoresis map of
Streptococcus pyogenes proteins. Unpublished data.; Best
Blastp hit = pir|A69581|acetyl-CoA carboxylase (biotin
carboxylase subunit) accC - Bacillus subtilis
>gi|1303913|dbj|BAAL2569.1| (D84432) ybhX [Bacillus
subtilis] >gi|2634868|emb|CAB14365.1| (Z99116) acetyl-CoA
carboxylase subunit (biotin carboxylase subunit) [Bacillus
subtilis]"
/codon_start=1
/transl_table=11
/product="putative acetyl-CoA carboxylase biotin
carboxylase subunit"
/protein_id="AAK34489.1"
/db_xref="GI:13622800"
/translations="MFKKILIANRGEIAVRIIRAAARELGISTVAVYSEADKEALHTIL
ADRAICIGARSRESYLNNSVLSAAIVTGAQAIHPGFLSENSKFATMCEEMNKF
IGPSASVMDKMGKINARSEMIAKAGVPVPGSDGVYNAQEAALANKIGVPMUKAS
AGGGGKIRKVEADLEAFNAASCEALGAFNGAMYLEKVIYPARHTEVQILDAY
GNIIHIGERDCSILORNNQVLRPEPSIAGNTLRLHMGQAARAAEAVANAGTIEF
LLDESEKVFEMNTRIQUERHPVFTGVDIVKEQIKAAQGLANQEDITIGH
ALBCRNAENTAFNPFSFGKITDLYMPGSGVGLRVDVSAVNGYAIPIPYDSMIAKII
VHGSNFDALMKRQALVELEIEGITITDQDLDTLSKRKRVAGDYDTSFLMETFLPH
YNKD"
/complement(3180..3599)
/gene="fabZ"
/note="SPY1746"

```

```

CDS
/note="Best Blastp hit = dbj|BAB16042.1| (AB030883)
similar to Bacillus subtilis hydroxymyristoyl-(acyl
carrier protein) dehydratase [Streptococcus
zooepidemicus]"
/codon_start=1
/transl_table=11
/product="putative beta-hydroxyacyl-ACP dehydratase"
/protein_id="AAK34490.1"
/db_xref="GI:13622801"
/translation="MDIREIQALPHRYPMLLVDRVLEVSDDHIVAKNVTINEFFFN
GHFPIKPMQPLVMEALQTAGVLELSKEENKGLVYAGMDKVKFKQVVPDGLV
MTAFPIKRGITIAVVEARAEDVGLAASGTUTPACGQ"
/complement(3599..4099)
/gene="accB"
/note="SPY1747"
/complement(3599..4099)
/gene="accB"
/note="Best Blastp hit = pir||C75558|acetyl-CoA
carboxylase, biotin carboxyl carrier protein - Deinococcus
radiodurans (strain R1)
>gi|6457783|gb|AAF09708.1|AE001874.5 (AE001874) acetyl-CoA
carboxylase, biotin carboxyl carrier protein [Deinococcus
radiodurans]"
/codon_start=1
/transl_table=11
/product="putative biotin carboxyl carrier protein"
/protein_id="AAK34491.1"
/db_xref="GI:13622802"
/translation="MNIQEKLDLNAQFDTSLSRFLFKTNEGELIFSKNEOHLNASTS
NOEHAVPVPQVLVPNTASEASPAVKDVFEEQPAESFVAGSDIVESPLVGVAY
LAASPDKPPFVAVGDTVKKGQTLVIIEMKVMNEVPAPCDGVTEILVSNEDVIEFGQ
GLVRIK"
/complement(4101..5333)
/gene="fabF"
/note="SPY1748"
/complement(4101..5333)
/gene="fabF"
/EC_number="2.3.1.41"
/note="The N-terminal amino acid sequence of this ORF has
been determined from a spot isolated by 2-D gel
electrophoresis from another strain of S. pyogenes. Hogan,
D. A., Whitton, M. M., Rogers, J. and R. A. VanBogelen.
2000. Two-dimensional gel electrophoresis map of
Streptococcus pyogenes proteins. Unpublished data.; Best
Blastp hit = pir||G69842|3-oxoacyl-(acyl-carrier-protein)
synthase [EC 2.3.1.41] - Bacillus subtilis
>gi|2633471|emb|CAB12975.1| (Z99109) similar to 3-oxoacyl-
acyl-carrier protein synthase [Bacillus subtilis]
>gi|2633488|emb|CAB12991.1| (Z99110) similar to 3-oxoacyl-
acyl-carrier protein synthase [Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="putative beta-ketoacyl-ACP synthase II"
/protein_id="AAK34492.1"
/db_xref="GI:13622803"
/translations="MTEKRVVVTGYGLTSPIDHDPETFWNLKAGQIGIPITKFDTT
DIAVNAAEIQDFPFKFKKQDLNFRDLSLYALAAKBAIINHADINIMVSDRFRG
VIVASGGIAIEIEQVIRLHEKGPVKPTLPKALPNWAGNVAIQAQGVCKSI
NTACASNDIAGDAFRAIKFTGQDVMIVGSEAAITKFAIQAQSLTALSTDPERS
SIPDKDRNGFIMGEGSLVLESLHAQERGAITLAEIVGYNTCDAYHMTSPNPEG
LGAKATHALQAEATASAINVNAHGTSTPANKEGSOAIVAVLGKDVPSVSKSF
TGHLGAAGAIEATATIEAVRHNVVPTAGTQALSEIEIANVIFGEGKETAINVAISN
TFGFGHNAVLAKCEE"
/complement(5348..6082)
/gene="fabG"
/note="SPY1749"
/complement(5348..6082)
/gene="fabG"
/EC_number="1.1.1.100"
/note="The N-terminal amino acid sequence of this ORF has
been determined from a spot isolated by 2-D gel

```

electrophoresis from another strain of *S. pyogenes*. Hogan, D. A., Whitton, M. M., Rogers, J. and R. A. VanBogelen. 2000. Two-dimensional gel electrophoresis map of

Streptococcus pyogenes proteins. Unpublished data.; Best Blast hit = dbj|BAB06210.1| (AP001515)

3-oxoacyl-(acyl-carrier protein) reductase [Bacillus

halodurans]

/codon_start=1

/transl_table=11

/product="putative beta-ketoacyl-ACP reductase"

/protein_id="AAK34493.1"

/db_xref="GI:13622804"

/translation="MEIKGKIFITGSTRIGLAWHOFASLEANIVLNGRSALSEBEL

VASFTDYGVTVTTSSEASEAKWNEAESIGSDLVNNAGITNDKMLKYTE

EDPERVLKINTGARNTQSVLKPIMKARQALINVSIVGLTGNIGQANYASAKAGM

IGFTKSAREVAARNICVNAIPFISDMTGVLPKMQEQILSIQIPKIRIGKAQEYA

HLASFLVEQDITGVIAIDGGMNQ

complement (6072..7010)

47.7%; Score 439.4; DB 1; Length 10909;

Best Local Similarity 67.3%; Pred. No. 1.8e-109;

Matches 620; Conservative 0; Mismatches 301; Indels 0; Gaps 0;

QY 1 ATGACTAAACAGCGCTTTTATTCTGCTCAAGTGCCTCAGTATCTAGGATGGACGG 60

DB 7010 ATGACAAAGACAGCGCTTTTATTCTGCTCAAGTGCCTCAGTATCTAGGATGGACGG 6951

QY 61 GATTCTTATGATCAGTATCGGATCTCAAGAAAGATTGATCGAGCGAGTCAGTGTCT 120

DB 6950 GATTCTTATGATCAGTATCGGATCTCAAGAAAGATTGATCGAGCGAGTCAGTGTCT 6891

QY 121 GATTATGATTACGTTATCTATCATCGATACGGAAGAACAACTCAATCAGACCGGTAT 180

DB 6890 GGATACGATTTCGGCGGTTCATTGATGATGACGAGTTTAAACTTAATCAAACTAGCTAT 6831

QY 181 ACGCAACAGCGATCTAGCGACTTCGGTGTCTATCTACCGTTTATTCGAAAGAAAGCGC 240

DB 6830 ACGCAACAGCGATCTAGCGACTTCGGTGTCTATCTACCGTTTATTCGAAAGAAAGCGC 6771

QY 241 TATCAGCTGATATGTTGCTGTTGTCTCTTGGAGATACTCTGCTTGGTGGCAAGC 300

DB 6770 GTTAAACGGATATGTTGCTGTTGTCTCTTGGAGATACTCTGCTTGGTGGCAAGC 6711

QY 301 GCGCGCTTGGATTTGAGATCGGTTGCTTGGTGTAGCTAGCGTGGAGCTATATGGAA 360

DB 6710 GCGCGCTTGGATTTGAGATCGGTTGCTTGGTGTAGCTAGCGTGGAGCTATATGGAA 6651

QY 361 GAAGCGGCTCTGCTGCTCTGCGAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGTCT 420

DB 6650 GAAGCGGCTCTGCTGCTCTGCGAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGTCT 6591

QY 421 ATGGAAGAGCTGTCAAAAGCTTCTGAACTTGGAGTGGTTACTTCCAGCCAACTATAAC 480

DB 6590 ATGGAAGAGCTGTCAAAAGCTTCTGAACTTGGAGTGGTTACTTCCAGCCAACTATAAC 6531

QY 481 ACACCTGCAAAATCGTCAATCTCTGAGAGTGGTTGAGTGTGATCGAGCGGTTGACCTT 540

DB 6530 ATCTCTAGTCAAAATCGTCAATCTCTGAGAGTGGTTGAGTGTGATCGAGCGGTTGACCTT 6471

QY 541 TTGCAAGAGAGGTCGCAAAAGCTTGTATCTCTTAAAGGTGTCAAGTCCCTTTTCAACC 600

DB 6470 TTGCAAGAGAGGTCGCAAAAGCTTGTATCTCTTAAAGGTGTCAAGTCCCTTTTCAACC 6411

QY 601 GCTCTCTTGTAGCTGTAGCGAGAACTAGCTGAACTCTAGCTCAGTAACTTTTCA 660

DB 6410 GCTCTCTTGTAGCTGTAGCGAGAACTAGCTGAACTCTAGCTCAGTAACTTTTCA 6351

QY 661 GATTCTTACTTGTCCCTAGTCGCGCAATACAGAGCTGTGTGATGCAAAAAGAGGACATT 720

DB 6350 GACTTCAAGATTCCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 6291

QY 721 GCTCAGCTCTTGAAGCGGTCAGTCAAGAAACCGTCTGTTCTATGAAGTATTTGGGTC 780

RESULT 10

AE014165/c

LOCUS

DEFINITION

50416 bp DNA linear BCT 19-JUL-2002

Streptococcus pyogenes MGAS315, section 30 of the complete

genome.

AE014165

AE014165.1

GI:21905237

ACCESSION

AE014165

VERSION

AE014074

KEYWORDS

Streptococcus

ORGANISM

Streptococcus pyogenes MGAS315.

Streptococcus pyogenes MGAS315

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Streptococcus.

REFERENCE

1 (bases 1 to 50416)

Beres, S.B., Sylva, G.L., Barbican, K.D., Lei, B., Hoff, J.S.,

Mammarella, N.D., Liu, M.-Y., Smoot, J.C., Porcella, S.F.,

Parkins, L.D., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M. and

Musser, J.M.

Genome sequence of a serotype M3 strain of group A *Streptococcus*:

phage-encoded toxins, the high-virulence phenotype, and clone

emergence

Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10078-10083 (2002)

JOURNAL

REFERENCE

2 (bases 1 to 50416)

Beres, S.B., Sylva, G.L., Barbican, K.D., Lei, B., Hoff, J.S.,

Mammarella, N.D., Liu, M.-Y., Smoot, J.C., Porcella, S.F.,

Parkins, L.D., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M. and

Musser, J.M.

Direct Submission

Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,

Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,

Hamilton, MT 59840, USA

FEATURES

Location/Qualifiers

1..50416

/organism="Streptococcus pyogenes MGAS315"

/strain="MGAS315"

/serotype="M3"

/db_xref="taxon:199466"

/note="group: A"

117..1103

/gene="SpvM3_1493"

117..1103

/gene="SpvM3_1493"

/note="best Elastp hit: gb|BAAL98256.1| (AB010083) putative

esterase [Streptococcus pyogenes MGAS8232]"

/codon_start=1

/transl_table=11

/product="putative esterase"

/protein_id="AAM80100.1"

/db_xref="GI:21905238"

/translation="MKHPRTKTLVTLGLLLTLCLPTAAASRSKSWFIEQYFWLKR

DKSYKSDQDFSFQYLDACREQSPYQLODNLVGLVQENLVCMQVSNWNGKP

LDKTYIAGGYSILNNPTTYHINMLKTSLSLAKIVLFIYPKAPRYINYNPKLVN

LYQHYKQNVFLMGDSAGGLALGLAHNLNVSVPQKLVLSPLDVTMSHPEI

PYEDADPTLSSGLKRVGLWAYSDNTNHIYVSPKNGPITLPLTFTGTREIFY

PIQRYAAKLKAANHTFITQEGNHHVPIPIEAKTAQOIIDAINKT"

complement (1239..1589)

/gene="rbfa"

/note="SpvM3_1494"

complement (1239..1589)

CDS

gene	<p> <code>/gene="rbfa"</code> <code>/note="best blastp hit: gb AAK34465.1 (AE006601) putative</code> <code>ribosome binding factor A [Streptococcus pyogenes M1 GAS]"</code> <code>/codon_start=1</code> <code>/transl_table=1</code> <code>/product="putative ribosome binding factor A"</code> <code>/protein_id="AAM80101.1"</code> <code>/db_xref="GI:21905239"</code> <code>/translation="MANHRRDQVMEIKREVMNIILOKKVRDPVQGVTTIEVMOQDGL</code> <code>SLAKVYTIIMSDLASDNKQADQGLGKATGTIKRELGQLTMYKIPDLVFEKDSIAIYAG</code> <code>NKIDQLRLDLDNKS"</code> <code>complement (1795. .4599)</code> <code>/gene="infB"</code> <code>/note="SpyM3_1495"</code> <code>complement (1795. .4599)</code> </p>
CDS	<p> <code>/gene="infB"</code> <code>/note="best blastp hit: gb AAK34466.1 (AE006601) putative</code> <code>initiation factor 2 [Streptococcus pyogenes M1 GAS]"</code> <code>/codon_start=1</code> <code>/transl_table=1</code> <code>/product="putative initiation factor 2"</code> <code>/protein_id="AAM80102.1"</code> <code>/db_xref="GI:21905240"</code> <code>/translation="MVSHAKYGLDVKSHASSVEEADAKKIIISFSKAKSPDVTASQT</code> <code>VKPEVRGVPNNHQTNEQDKRITTHRSQGTNDKRTERRKASNVSPHDHNLQVGRDN</code> <code>RSFAKHNELGRFTNOKKQGRQEPQSKPKIDFKAPAAALKAEQNAEYSRQSTPERA</code> <code>QQAERKALHARQAKAEALKAQAEEMSHREAAALKSTIEEAETLKSSNLSAKTADNR</code> <code>KKQARPEKNRHLTHSQEQGQKKNKSNQVRQKNSNNWNNKTKTKKGNKAKNTN</code> <code>TAPVTERKPEHLPKFEYTGEDTVAEIAKRIKREPAEIVKXLFPMGVNATQNSGLD</code> <code>GDTTELMVDYGEISAKKVEEDADIEFFEDENLPENIPERAPVTTIGMVRHGRK</code> <code>TTLDLIRNSVATGEAGGITQHIGAYQIEAGKKITFLDTPGHAFTSMRGARAGVIT</code> <code>DITILIVAADQVMPTQETIAHNSKAGVPIIVAINKIDKPGANPRVIAELAEVGI</code> <code>STANGGECEFVEISAKENKIDDELLETVLLVAEVELKADPTVRAITGVTEAPLDKKG</code> <code>CALATLVQOQTALVDPPIVVGNTFGVRVNRNDLGRVKVSAEPTSPVSIITGLNETPM</code> <code>AGDPVADSKARARQKNTQRVSLDNLDTLKAGEIKTNNVLIK</code> <code>ADVQSEALASALVKEIVGVNVVHSAVGNESDVTFLAESNAVITIGNVRPTP</code> <code>QARQADDTVEILHSIIYKVIIVEFEAMKGLDPVYQEKILGEAIIRTFPKVSKVG</code> <code>TIGGFMVINGKVTSDSVRVIRDSWVIFDGLKSLAKHYKDDVKEVNGAQEGGLMIENF</code> <code>NDLKVDOTIEAYIMEETVRK"</code> <code>complement (4676. .4978)</code> <code>/gene="SpyM3_1496"</code> <code>complement (4676. .4978)</code> </p>
gene	<p> <code>/gene="best Blastp hit: gb AAL98259.1 (AE010083) putative</code> <code>ribosomal protein [Streptococcus pyogenes MGAS8232]"</code> <code>/codon_start=1</code> <code>/transl_table=1</code> <code>/product="putative ribosomal protein"</code> <code>/protein_id="AAM80103.1"</code> <code>/db_xref="GI:21905241"</code> <code>/translation="MTNLERLSSLIQLAQRAGKVISGEELVVRKAIHQQQVILVFAND</code> <code>AGPNVTKVTDPKSNVNVSTVNLALSELGAALGPKRKVAAIADAGFSKKMRLME"</code> <code>complement (4971. .5267)</code> <code>/gene="SpyM3_1497"</code> <code>complement (4971. .5267)</code> </p>
CDS	<p> <code>/gene="SpyM3_1497"</code> <code>/note="best non-GAS blastp hit: gb AAK74710.1 (AE007365)</code> <code>conserved hypothetical protein [Streptococcus pneumoniae</code> <code>TIGR4], and gb AAK93283.1 (AE008427) Conserved</code> <code>hypothetical protein [Streptococcus pneumoniae R6]"</code> <code>/codon_start=1</code> <code>/transl_table=1</code> <code>/product="conserved hypothetical protein"</code> <code>/protein_id="AAM80104.1"</code> <code>/db_xref="GI:21905242"</code> <code>/translation="MSWKVPIPKSLVSGEIIAKRDLIRIVTKGQGVFIQDTFGKGN</code> <code>GRGAVIKLDNQEALMAKKKQVNEFSFMDIPESFYDDLIAYVDHKIKRRELGLD"</code> <code>complement (5283. .6428)</code> <code>/gene="rusA"</code> <code>/note="SpyM3_1498"</code> <code>complement (5283. .6428)</code> <code>/gene="rusA"</code> </p>
gene	
CDS	

/note="best blastp hit: gb|AAK34469.1| (AE006601)
transcription termination-antitermination factor
[Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative transcription
termination-antitermination factor"
/protein_id="AAK80105.1"
/db_xref="GI:21905243"
/translation="MUEAFRLSEEHKADIIIDAVTESLKAYKRRYQSSSCVIE
FNEKADQFTVREVEVEVDSRLSISUKDALISAYVELDGDKIRFESVNFGRVGA
AQAQOTIMKQRQREVMFNFSQHEGEIMTGTVERDQRFIYVNLGSLAQQLSHQ
DQIPGADITGRIVVYVKNVNNPKGVNVSRSHPETIKRMEGEIVFVDTGTVLPI
MYSRGAGDRTKVAISRSHNNVDIAIGTIVRGSGNNIKVVISKPHKPRVDAKTGLEIPV
BENIDVQWDDPAERTFYNAIAPEAVDMVLFDDEDLKRAIVVVPDSKLSLAIGRRQCN
VRLAHLHTGVRIIDIKASAEVDRLAEAKAATAVEEPVWDDATAEEV"
/complement(6615..7151)
/gene="SpyM3_1499"
/complement(6615..7151)
/gene="SpyM3_1499"
/note="best non-CAS blastp hit: gb|AAK99281.1| (AE008427)
Conserved hypothetical protein [Streptococcus pneumoniae
R6]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK80106.1"
/db_xref="GI:21905244"
/translation="MDSQGIPILEKIKIEBVIKIANTSIIDITVTKTPEIKAPYEL
VDVEYKMGSDYILSLIVDKEDGITVEDTDLNIIISPLDITDPPFPNQMYLVESS
PGLERPLKTAADSLKAAVGSYINSLYQAIDKKVQFGDLLAFDGETLITDYLDKTRHK
IVNIPLYQAVAKVRMAVKL"
/complement(7259..7345)
/product="crna-ser"
/complement(7410..8045)
/gene="SpyM3_1500"
/complement(7410..8045)
/gene="SpyM3_1500"
/note="best Blastp hit: gb|AAL98263.1| (AE010083) putative
methyltransferase [Streptococcus pyogenes MGAS8232]"
/codon_start=1
/transl_table=11
/product="putative methyltransferase"
/protein_id="AAK80107.1"
/db_xref="GI:21905245"
/translation="MRVRKEKGAEEHLANNPHYVILNPEDAKGRHDFGNDRPIRIE
VGSKGKGFITGMALQPDINIGYIDIGLSVLSDALKLASVEVNVKLLRVDGSSLTN
YFDGEYDMVYALNFSQDPWPKTEKREKRLTYKFDLYKRIIPHGHSIFHFTDNRGLFE
YSLASFQYQWTLRQRLWDLHASNYSYGNWVTEYEKESNKGVIVRVEANF"
/complement(8045..8836)
/gene="SpyM3_1501"
/complement(8045..8836)
/gene="SpyM3_1501"
/note="best Blastp hit: gb|AAK99279.1| (AE008427)
Conserved hypothetical protein [Streptococcus pneumoniae
R6]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK80108.1"
/db_xref="GI:21905246"
/translation="MTTEQELTLTLPRKSGKAYKGYTNGECVFKLNTTLPILPAL
AKEQIAPOLLKAPYKMGNDMSQAQEWLNGRTLKEDMNSKQIILHLLRHKSKLVNQ
LLQNYAIEPNWLDVDFEQNAFLQCSYLQAIKVELKRSLEPFKSEVATIVHGDI
KHSNWVITTSGMILFLVDWDSVRLTDWYDVAVLLSHVIPSRSWSELSYGYKNNKDV
MQKLIWYQGSQSHLTQILKCDPDKDMEHVQGEIYALRKPFREIFRKK"
/complement(8900..9934)
/gene="SpyM3_1502"
/complement(8900..9934)
/gene="SpyM3_1502"
/note="best Blastp hit: gb|AAK34473.1| (AE006601) putative

Query Match

RESULT 11	AE010088/c	AE010088	14456 bp	DNA	linear	BCT 03-APR-2002
LOCUS		Streptococcus pyogenes strain MGAS232,				
DEFINITION		complete genome.				section 136 of 173 of the

```

/ gene="spyM18_1811"
2350..3159
/ gene="spyM18_1811"
/ note="best blastp match gb|AAK34483.1| (AE006602)
putative mannose-specific phosphotransferase system
component IID [Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative mannose-specific phosphotransferase
system component IID"
/ protein_id="AAL98331.1"
/ db_xref="GI:19748920"
/ translation="MSDISIISAILVVIIVIIAFFAGLEIGLIDFQFOMHPLVACTLTIGLVT
GHLEAGVILGGLTQMLALGANIAGVAPDAALASVAAAIIMVKSGDFQKGTTRAYS
TAIVLAVAGFLMTVITLTALVHAGDKAAGNFGAGIERPHEFIALLIGQRIAPVA
ALLVAVPTSAVQSVLANMPKMLNEHQIGAGVAVGVANVINMATTREVPWFALGF
ALAAISLTLIAGVIGVAFIYLNLSKGGNGNAAGSADFIGDILEDY"
3176..4087
/ gene="spyM18_1812"
3176..4087
/ gene="spyM18_1812"
/ note="best blastp match gb|AAK34484.1| (AE006602)
putative mannose-specific phosphotransferase system
component IID [Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative mannose-specific phosphotransferase
system component IID"
/ protein_id="AAL98332.1"
/ db_xref="GI:19748921"
/ translation="MTQIKLTKSDQRVWVRSQFLQGSNNYERMNQMGVAALIPAL
KMLYSPEDRAALERMEFNTHPYVAAPIIGVTLALBEERANGFPIDKAIQGVKI
GKMLPLAGIDGPVFWFTVRPILGALGASLASTGNIVGPLLFFFGWNLIRMAFLMYTQE
FYKAGSEITKMSGLQIDITKGAISLGMFILAIVQVRWVSINFIDILPGKQLSDCA
VYVPPDCAVGAEIKTILANAIGMSLIDKVAQTLQGLDLSLIPGLAGLLIFLGNWL
LKKVSPRIIIGLFAFGILAHAGIM"
4201..4560
/ gene="spyM18_1813"
4201..4560
/ gene="spyM18_1813"
/ note="best blastp match gb|AAK34485.1| (AE006602)
conserved hypothetical protein [Streptococcus pyogenes M1
GAS]"
/ codon_start=1
/ transi_table=11
/ product="conserved hypothetical protein"
/ protein_id="AAL98333.1"
/ db_xref="GI:19748922"
/ translation="MAQSLNTSVEYKTKAVSYLGMGKGVGHILLGPKALEFYNDKNVN
DYIQIPHTAINHIGANVSRRKVSRRHFEIFTDQKFLFASGDSKILKITRQHIGNEKV
ITLPTLMQITINKFFRH"
4952..6229
/ gene="serS"
4952..6229
/ gene="serS"
/ note="best blastp match gb|AAK34486.1| (AE006602)
putative seryl-tRNA synthetase [Streptococcus pyogenes M1
GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative seryl-tRNA synthetase"
/ protein_id="AAL98334.1"
/ db_xref="GI:19748923"
/ translation="MIDLKRIKRTDFTVAAKLKNRVSSEDTLTHLKELEKERTLLVQ
SEELKARNASAAIAQAQRKEDATQQIADMOKVSADIKTIDNQLVAIDQVADIIIT
VLPTNPHDSVEFGADENVEIRWGTFRDFEVKAHMDLGEDLDILDMRGKAVYTG
ARFLFYKXNLGALERALYNPMLDEHKEGVEIITPYMVNHDMSMFGTGOYVPKFKEDTF
ELADTNFLVLPFAEVPYNYVGEILLDKGLPIYFTAMGPSRPSRAGSAGRTDRLIR
LHOFHKVMYKFKPESYOELEKMTANAENILKLGIPYVLSLCTGMPGSRKATY
DLWVIPAQNTYREIISCSNTEDFQARRAQIRYRDEADGKVLHLLHNLNGLAVGRVY
AAILNQNDGDSVTIFEPFMGETVISPX"
complement (6462..7232)

```

```

/ gene="acca"
/ note="spyM18_1815"
complement (6462..7232)
/ gene="acca"
/ note="best blastp match gb|AAK34487.1| (AE006603)
putative acetyl-CoA carboxylase alpha subunit
[Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative acetyl-CoA carboxylase alpha subunit"
/ protein_id="AAL98335.1"
/ db_xref="GI:19748924"
/ translation="MTDVSRYLKEARDQGRITLTDYANLIFFDFMELHGRHFSDDGA
IVGSLAVLAGPVPVIGIKQKNLQHNLRNFGQPNPEGVKALRLKQAKRFRPVPV
TFINTAGYFVGAEKGGGGAIAKNLMMNSDLKVPITAIILIGGSGGALALAVADQ
VMMLNTWYAVLSPEGFASILWKQDSRATAEALMKKITAGELYQMGIVDRIIPBHGYP
SSEIVDIIRKANLIBQITSLQAKPLDQLDLDERYQRFKY"
complement (7229..8095)
/ gene="accD"
complement (7229..8095)
/ note="spyM18_1816"
complement (7229..8095)
/ gene="accD"
/ note="best blastp match gb|AAK34488.1| (AE006603)
putative acetyl-CoA carboxylase beta subunit
[Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative acetyl-CoA carboxylase beta subunit"
/ protein_id="AAL98336.1"
/ db_xref="GI:19748925"
/ translation="MALFRKDKXIRITPNNFLKGSVSHNVPEVPDELFAKCPACKHM
IYKQDLGAKICTPCTSNFRISAQERLTLTVDEGSFOELFTSIETKQPLRFPYQEKI
OKKDETLGLHVAITGKAMVQGOIALAIDMSHFIMASMGTVGKELTRLELAIEEN
LPVITFASGARMQEGIMSLQMAKVSAAVKRHSNAGLFYLTITLPTTGTGVTASPA
MEGDIILAEPSQLVGFAGSERVETVRENLPDDFKAEFLQDHGFVDAIVKRELRDK
IAHLVAFRGGGQ"
complement (8104..9468)
/ gene="accC"
complement (8104..9468)
/ note="spyM18_1817"
complement (8104..9468)
/ gene="accC"
/ note="best blastp match gb|AAK34489.1| (AE006603)
putative acetyl-CoA carboxylase biotin carboxylase subunit
[Streptococcus pyogenes M1 GAS]"
/ codon_start=1
/ transi_table=11
/ product="putative acetyl-CoA carboxylase biotin
carboxylase subunit"
/ protein_id="AAL98337.1"
/ db_xref="GI:19748926"
/ translation="MFKKILIANGEIAVRIIRAARELGISTVAVYSEADKEALHTLIL

```

```

Query Match      47.5%; Score 437.8; DB 1; Length 14456;
Best Local Similarity 67.2%; Pred. No. 4.9e-109;
Matches 619; Conservative 0; Mismatches 302; Indels 0; Gaps 0;

QY 1 ATGACTAAACAGCCCTTTTATTTGCTGCTCAAGTGCCCGCAGTATCTAGGATGGACGG 60
DB 13330 ATGACAAAACAGCTTTTATTTGTCGCGTGCAGGTGCCAAAAATTTAGGATGCCAAG 13271
QY 61 GATTTCTATGATCAGTATCCGATTGTCAAGAAACGATTGATCGAGCGAGTCAGTGCTC 120
DB 13270 GATTTTATGATTACTTTTGCTATTGTAGAAAAAACTTTTGATCAAGCTAGTCAAGTATTG 13211
QY 121 GGTATGATTTACGTTATCTCATCGATACGGAAGAACAACTCAATCAGACCCGCTAT 180
DB 13210 GGATACGATTTGCGCGCTTTGATTGATAGTACGAGTATAAACTTAATCAAACTAGCTAT 13151
QY 181 ACGCAACACAGCATTCTAGCAGCTTCGGTTCCTATCTACCGTTTATTTGCAAGAAAGGCG 240
DB 13150 ACGCAACAGCTATTCTTGCATCATCAATTCCTATTACCGGTTTGGGGCTTACATCAC 13091
QY 241 TATCAGCCTGATATGTTGCTGTTGTTCTCTTGGAGAACTACTCTGCTTGGTGGCAAGC 300

```



```

/translation="MTRNVISGVYVSAIGNSPEEFNNLKAGKTGIGPIHFDARA
TGISVAEVEFEDKQKADPRMDTFSIYAVYALDAMWMSGITEENTHEDRLGC
IMAGIGLESGQNSQAIYAKGRKGVAGLYVPLAIANWATGNVALRTGARGVSRAE
VTCAGANSIGSAREIKGIDAMTAGGAAICELGAGFANULTAFKEDTPPEKA
CIPDKRSGVMEGGGLVLESLEHAQAGANILAEIVGYNTDNALFTPTPSGVG
AENAKI LAEAGASPADYVNAHGTSTHANEETSKAHNVIGDKAYVSSYKALHG
HALGATIAEAVATVQALLNQYAPVNAAGTTELDGMTINVVLGEGRKADINLALSQDF
GFGGHNVAFAFKWTGE"
gene
3635..4102
/ gene="accB"
/ note="LOC8"
3635..4102
/ gene="accB"
/ note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/ codon_start=1
/ trans_table=11
/ product="biotin carboxyl carrier protein of acetyl-CoA
carboxylase (EC 6.4.1.2)"
/ protein_id="AAK04874.1"
/ db_xref="GI:12723695"
/ translation="MNISEVKDLMNQFNGSTLREFSKNADGELSPSKNGOALVGAQ
SAAPTAPFAQAQTPVAESNEAPVASEAAVEAGEAVTSPLVGAVYIKPSPKAEFA
KYGSVVKGTGLLIEAMKWNNEIPAPKGVITHEINVSGEDVIEFGODLMRIK"
4229..4663
/ gene="fab22"
/ note="LOC188"
4229..4663
/ gene="fab22"
/ note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/ codon_start=1
/ trans_table=11
/ product="3R-hydroxymyristoyl-(3R)-HYDROXYMYRISTOYL-acyl
carrier protein dehydratase (EC 4.2.1.1)"
/ protein_id="AAK04875.1"
/ db_xref="GI:12723696"
/ translation="MTEVNVINVEIMEALPHRYPPFLVDRVIDIAEDBITAKNVTIN
EEFFQGHFPQIPVMGVLIMEAQAAGVLELSKPNKGLVFIAGMDNVKRYKKQVTP
GDQVLVHAKFIKRGPIAVVEAATVDGKLAAXGTLTFALGR"
5057..6424
/ gene="accC"
/ note="LOC189"
5057..6424
/ gene="accC"
/ note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/ codon_start=1
/ trans_table=11
/ product="biotin carboxylase (EC 6.3.4.14)"
/ protein_id="AAK04876.1"
/ db_xref="GI:12723697"
/ translation="MFNKKILIANRGEIAVRIIRAARELGVGTVAIYSTADRDALHVLH
ADEALCVGPASTESYLMNNILSAAVTAQAIHPGFGFLSENSFARLCEBMINIKF
IPGAKVMKKNVNEELVPAFEAAQAAQAFNGAMFIERMIFPAHLEVOILGDSQ
AGGKGKIRKVVNVEELVPAFEAAQAAQAFNGAMFIERMIFPAHLEVOILGDSQ
GNVHLGERCSLQNNQVLESFSAIGTILRNEITAAVAZAAQHVSYENAGTIEP
LLEDSKGFIFMNEVQVHEVPFVSGVDIVKEQIRVANGNELSVKQEDVTFSGH
AIECRINAEPKNFAPSPGKINLFLPSGGVGLRVDMSAYSGISIPPYDSDMIKVI
VHGENRFEALMKMORALEFDEVGVIITNDVQLELISDPHVHVGADYDTSFLGNVFLPE
YQKED"
6510..7376
/ gene="accD"
/ note="LOC180"
6510..7376
/ gene="accD"
/ note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FATTY ACID BIOSYNTHESIS"
/ codon_start=1
/ trans_table=11
/ product="acetyl-CoA carboxylase carboxyl transferase

```

```

subunit beta (EC 6.4.1.2)"
/ protein_id="AAK04877.1"
/ db_xref="GI:12723698"
/ translation="MALFKKKYIKINPNRSIIIEKQARQEPVDPDLFAKCPACKHTIY
QKDLGKVCNCDYNFRITAKERLAIYADKDSFVEMFTGIESKNPLDFPGYPEKLA
TKARGLDEAVITGATIKGQKTALAIINDSTFINASMGTVVGEKILTRIFEVATPKLP
IIVFTASGGARMEQIGIMSLMQMAKTSAAVKRHSNAGLFYITVLTPTTGGVTASPASL
GDIIILAEPSQLIGFAGRRVISOVTRQLPDDFQKAEFLNLHGFDVAIVKRTRELQKLA
LLELHTEVENV"
7369..8151
/ gene="accA"
/ note="LOC190"
7369..8151
/ gene="accA"
/ note="EVIDENCE BY HOMOLOG Y BIO07.01 FATTY ACID AND
PHOSPHOLIPID METABOLISM. General. FIRST STEP IN LONG-CHAIN
FATTY ACID SYNTHESIS"
/ codon_start=1
/ trans_table=11
/ product="acetyl-CoA carboxylase carboxyl transferase
subunit alpha (EC 6.4.1.2)"
/ protein_id="AAK04878.1"
/ db_xref="GI:12723699"
/ translation="MSEIIDSILTGEVSQKNIADIINEARAADRISARDIILDKVFTDFE
ELHGRQKTDDEALVGGIARFNQAVTIGIKGRNLAEINLATNFGSQSPNGRYKALR
LMQKREKFRPVTIFITAGAPGIEAEERGCGEAIARNLLEMSDLKPIIAFIATFEGE
GSGGALALANKVYILENAVSVLSPEGFATILWKDGSRRDEAAELMKITAPHLLEM
GIVDAIISEENLKRQLTESLGLSLLTETELAEQRYERFRKY"
8280..9422
/ gene="metB2"
/ note="LOC191"
Query Match 42.2%; Score 388.8; DB 1; Length 10701;
Best Local Similarity 64.6%; Pred. No. 1.4e-95;
Matches 598; Conservative 0; Mismatches 322; Indels 6; Gaps 1;
QY 1 ATGACTAAACAGCGCTTTTATTCGTGTCGTCAGGTGCCAGTATCTAGGATGGACGG 60
DB 615 ATGACTAAACAGCAATTTTATTCAGTCAAGGGCACAAAGCTTGGAAATGCACGT 674
QY 61 GATTTCTATGATCAGTATCCGATTCGTCAGGAAACGATTTGATCGAGCGAGTCAGGTGCT 120
DB 675 GACTTATATGACCAATATGAAACAGTTAAAGCAACTTTTGTATGAAGCAAGTCAAGCTTTA 734
QY 121 GCTTATGATTTACGTATCTCATCGATACGGAGAGACAACTCAATCAGACCCGCTAT 180
DB 735 GGATATGATTTTCGAGCTTTGATTAATGATGAAGAAAAACTAAATGAACAAAGTAC 794
QY 181 AGCCAACCAAGCCATTTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240
DB 795 ACTCAACCTGCAATTTTAAACACCTCTGTGCTATTTTACGTTTGTAAAGTGAATGGG 854
QY 241 TATCAGCCTGATATGGTTGCTGTTTGTCTCTTGAGAAATACTCTGCCTTGGTGCAAGC 300
DB 855 ATTAAACCTGACCTTGTTCGTGCTTGTAGTCTTGTGTAATATTCTGCCTTGGTATCATCA 914
QY 301 GCGCGCTTGGATTTTGAAGATGCGGTTGCCCTTGGTAGCTAAAGCGTGGAGCCTATATGAA 360
DB 915 GGAATCATTTGATTTTCAAGAGACAGTTAAGCTTTAGCTAACGTTGGTCAATATATGACA 974
QY 361 GAAGCGGCTCTGCTGATCTCGGCAAGATGTTAGCAGTTCTCAATACGCCAGTAGAGATC 420
DB 975 GAAGCTGACCCAGCTGTTCTGTTAAATGGTTCGCTTAATGAACACAGACCCAGCCTTG 1034
QY 421 ATTGAAGAGCGCTGTCAAAAAGCTTCTGAACCTT-----GGAGTGGTTACTCCAGCCCAAC 474
DB 1035 ATTGAAGAAATTTGCCAAGAGCGCCGAAATTAAGCGCGGTATTGTAGTCCAGCAAT 1094
QY 475 TATAACACACCTTGCAAAATCGTCAATTCGTGAGAAAGTGGTTGCAAGTTGATCGAGCGGTT 534
DB 1095 TATAACACGCCCGCACAAATTTGTTATGTTGTTGAAGTTGAGCGGTTGATTATCTGCTTT 1154
QY 535 GAACTTTTGCAAGAGCAGGTGTCACCAACGCTTGATTCCTCTTAAAGTGTCTAGGTCCTCTTT 594

```

```
Db 1155 GAGTGTCTAAAGAGCCGAGCTTCGTAACCTATTGTAATTAAGTTTCAGGACCTTC 1214
QY 595 CACACCGCTCTCTTGGAGCTCTAGCCAGAACTAGCTGAACTCTAGCTCAGTCAAGT 654
Db 1215 CATACAGCAATTTAAACACGACATCTGAAAATTTGGCTTGGAGCTTGATATAATGAT 1274
QY 655 TTTTCAGATTTACTTGTCCCTAGTCGGCAATACAGAACTGCTGCTGATGCAAAAAGAG 714
Db 1275 TTTAGACCTTTTGAATTACCATTAATCTCAATACAGAGTCTGCTGAAAGTAAAGAAATGAT 1334
QY 715 GACATTGCTCAGCTCTTACGCGCTCAGGTCAGGCAAGCCGTTCTGTTCTATGCAAGATT 774
Db 1335 GAAGTCAAGGACTTTTACGCGCTCAGGTCAGGCAAGCCGTTCTGTTCTATGCAAGATT 1394
QY 775 GGGGTCATCAAGAGCAGGCAATAGCAACTTTATCGAGATGGACCGGGAAGTCTTG 834
Db 1395 GAAACAATCAAAAACCTAGGCGGACTCGCTTTATGAATGGTCTCGGAGAGTACTT 1454
QY 835 TCAGGTTTGTAAAAAATTCATCAACTGCTCCTACTAGCTCATGTCGAGATCAGCG 894
Db 1455 TCAGGTTTCAATTAATAAATTCATCAACTGCTCCTACTAGCTCATGTCGAGATCAGCG 954
QY 895 AGTTTAGTACGACTTTTGAAGAAATA 920
Db 1515 TCATTTGAAGCTTGTGATTAATCAGTA 1540

RESULT 13
AE004276/c
LOCUS
DEFINITION
14530 bp DNA linear BCT 31-JUL-2000
Vibrio cholerae chromosome I, section 184 of 251 of the complete
chromosome.
ACCESSION AE004276
VERSION AE004276.1
KEYWORDS GI:9656555
SOURCE
Vibrio cholerae
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
REFERENCE
1 (bases 1 to 14530)
AUTHORS
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Baas, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C. and White, O.
TITLE
DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae
JOURNAL
Nature 406 (6795), 477-483 (2000)
MEDLINE
20468833
PUBMED
10952301
REFERENCE
2 (bases 1 to 14530)
AUTHORS
Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.A.,
Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
Ermolaeva, M.D., Vamathevan, J., Baas, S., Qin, H., Dragoi, I.,
Sellers, P., McDonald, L., Utterback, T., Fleischmann, R.D.,
Nierman, W.C., White, O., Salzberg, S.L., Smith, H.O., Colwell, R.R.,
Mekalanos, J.J., Venter, J.C. and Fraser, C.M.
TITLE
Direct Submission
JOURNAL
Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
FEATURES
source
1..14530
/organism="Vibrio cholerae"
/strain="N16961"
/serotype="O1"
/db_xref="taxon:666"
/chromosome="I"
/notes="biotype: El Tor"
223..1806
/gene="VC2012"
223..1806
/gene="VC2016"
223..1806
/gene="VC2018"
223..1806
/gene="VC2019"
223..1806
/gene="VC2020"
223..1806
/gene="VC2021"
223..1806
/gene="VC2022"
223..1806
/gene="VC2023"
223..1806
/gene="VC2024"
223..1806
/gene="VC2025"
223..1806
/gene="VC2026"
223..1806
/gene="VC2027"
223..1806
/gene="VC2028"
223..1806
/gene="VC2029"
223..1806
/gene="VC2030"
223..1806
/gene="VC2031"
223..1806
/gene="VC2032"
223..1806
/gene="VC2033"
223..1806
/gene="VC2034"
223..1806
/gene="VC2035"
223..1806
/gene="VC2036"
223..1806
/gene="VC2037"
223..1806
/gene="VC2038"
223..1806
/gene="VC2039"
223..1806
/gene="VC2040"
223..1806
/gene="VC2041"
223..1806
/gene="VC2042"
223..1806
/gene="VC2043"
223..1806
/gene="VC2044"
223..1806
/gene="VC2045"
223..1806
/gene="VC2046"
223..1806
/gene="VC2047"
223..1806
/gene="VC2048"
223..1806
/gene="VC2049"
223..1806
/gene="VC2050"
223..1806
/gene="VC2051"
223..1806
/gene="VC2052"
223..1806
/gene="VC2053"
223..1806
/gene="VC2054"
223..1806
/gene="VC2055"
223..1806
/gene="VC2056"
223..1806
/gene="VC2057"
223..1806
/gene="VC2058"
223..1806
/gene="VC2059"
223..1806
/gene="VC2060"
223..1806
/gene="VC2061"
223..1806
/gene="VC2062"
223..1806
/gene="VC2063"
223..1806
/gene="VC2064"
223..1806
/gene="VC2065"
223..1806
/gene="VC2066"
223..1806
/gene="VC2067"
223..1806
/gene="VC2068"
223..1806
/gene="VC2069"
223..1806
/gene="VC2070"
223..1806
/gene="VC2071"
223..1806
/gene="VC2072"
223..1806
/gene="VC2073"
223..1806
/gene="VC2074"
223..1806
/gene="VC2075"
223..1806
/gene="VC2076"
223..1806
/gene="VC2077"
223..1806
/gene="VC2078"
223..1806
/gene="VC2079"
223..1806
/gene="VC2080"
223..1806
/gene="VC2081"
223..1806
/gene="VC2082"
223..1806
/gene="VC2083"
223..1806
/gene="VC2084"
223..1806
/gene="VC2085"
223..1806
/gene="VC2086"
223..1806
/gene="VC2087"
223..1806
/gene="VC2088"
223..1806
/gene="VC2089"
223..1806
/gene="VC2090"
223..1806
/gene="VC2091"
223..1806
/gene="VC2092"
223..1806
/gene="VC2093"
223..1806
/gene="VC2094"
223..1806
/gene="VC2095"
223..1806
/gene="VC2096"
223..1806
/gene="VC2097"
223..1806
/gene="VC2098"
223..1806
/gene="VC2099"
223..1806
/gene="VC2100"
223..1806
/gene="VC2101"
223..1806
/gene="VC2102"
223..1806
/gene="VC2103"
223..1806
/gene="VC2104"
223..1806
/gene="VC2105"
223..1806
/gene="VC2106"
223..1806
/gene="VC2107"
223..1806
/gene="VC2108"
223..1806
/gene="VC2109"
223..1806
/gene="VC2110"
223..1806
/gene="VC2111"
223..1806
/gene="VC2112"
223..1806
/gene="VC2113"
223..1806
/gene="VC2114"
223..1806
/gene="VC2115"
223..1806
/gene="VC2116"
223..1806
/gene="VC2117"
223..1806
/gene="VC2118"
223..1806
/gene="VC2119"
223..1806
/gene="VC2120"
223..1806
/gene="VC2121"
223..1806
/gene="VC2122"
223..1806
/gene="VC2123"
223..1806
/gene="VC2124"
223..1806
/gene="VC2125"
223..1806
/gene="VC2126"
223..1806
/gene="VC2127"
223..1806
/gene="VC2128"
223..1806
/gene="VC2129"
223..1806
/gene="VC2130"
223..1806
/gene="VC2131"
223..1806
/gene="VC2132"
223..1806
/gene="VC2133"
223..1806
/gene="VC2134"
223..1806
/gene="VC2135"
223..1806
/gene="VC2136"
223..1806
/gene="VC2137"
223..1806
/gene="VC2138"
223..1806
/gene="VC2139"
223..1806
/gene="VC2140"
223..1806
/gene="VC2141"
223..1806
/gene="VC2142"
223..1806
/gene="VC2143"
223..1806
/gene="VC2144"
223..1806
/gene="VC2145"
223..1806
/gene="VC2146"
223..1806
/gene="VC2147"
223..1806
/gene="VC2148"
223..1806
/gene="VC2149"
223..1806
/gene="VC2150"
223..1806
/gene="VC2151"
223..1806
/gene="VC2152"
223..1806
/gene="VC2153"
223..1806
/gene="VC2154"
223..1806
/gene="VC2155"
223..1806
/gene="VC2156"
223..1806
/gene="VC2157"
223..1806
/gene="VC2158"
223..1806
/gene="VC2159"
223..1806
/gene="VC2160"
223..1806
/gene="VC2161"
223..1806
/gene="VC2162"
223..1806
/gene="VC2163"
223..1806
/gene="VC2164"
223..1806
/gene="VC2165"
223..1806
/gene="VC2166"
223..1806
/gene="VC2167"
223..1806
/gene="VC2168"
223..1806
/gene="VC2169"
223..1806
/gene="VC2170"
223..1806
/gene="VC2171"
223..1806
/gene="VC2172"
223..1806
/gene="VC2173"
223..1806
/gene="VC2174"
223..1806
/gene="VC2175"
223..1806
/gene="VC2176"
223..1806
/gene="VC2177"
223..1806
/gene="VC2178"
223..1806
/gene="VC2179"
223..1806
/gene="VC2180"
223..1806
/gene="VC2181"
223..1806
/gene="VC2182"
223..1806
/gene="VC2183"
223..1806
/gene="VC2184"
223..1806
/gene="VC2185"
223..1806
/gene="VC2186"
223..1806
/gene="VC2187"
223..1806
/gene="VC2188"
223..1806
/gene="VC2189"
223..1806
/gene="VC2190"
223..1806
/gene="VC2191"
223..1806
/gene="VC2192"
223..1806
/gene="VC2193"
223..1806
/gene="VC2194"
223..1806
/gene="VC2195"
223..1806
/gene="VC2196"
223..1806
/gene="VC2197"
223..1806
/gene="VC2198"
223..1806
/gene="VC2199"
223..1806
/gene="VC2200"
223..1806
/gene="VC2201"
223..1806
/gene="VC2202"
223..1806
/gene="VC2203"
223..1806
/gene="VC2204"
223..1806
/gene="VC2205"
223..1806
/gene="VC2206"
223..1806
/gene="VC2207"
223..1806
/gene="VC2208"
223..1806
/gene="VC2209"
223..1806
/gene="VC2210"
223..1806
/gene="VC2211"
223..1806
/gene="VC2212"
223..1806
/gene="VC2213"
223..1806
/gene="VC2214"
223..1806
/gene="VC2215"
223..1806
/gene="VC2216"
223..1806
/gene="VC2217"
223..1806
/gene="VC2218"
223..1806
/gene="VC2219"
223..1806
/gene="VC2220"
223..1806
/gene="VC2221"
223..1806
/gene="VC2222"
223..1806
/gene="VC2223"
223..1806
/gene="VC2224"
223..1806
/gene="VC2225"
223..1806
/gene="VC2226"
223..1806
/gene="VC2227"
223..1806
/gene="VC2228"
223..1806
/gene="VC2229"
223..1806
/gene="VC2230"
223..1806
/gene="VC2231"
223..1806
/gene="VC2232"
223..1806
/gene="VC2233"
223..1806
/gene="VC2234"
223..1806
/gene="VC2235"
223..1806
/gene="VC2236"
223..1806
/gene="VC2237"
223..1806
/gene="VC2238"
223..1806
/gene="VC2239"
223..1806
/gene="VC2240"
223..1806
/gene="VC2241"
223..1806
/gene="VC2242"
223..1806
/gene="VC2243"
223..1806
/gene="VC2244"
223..1806
/gene="VC2245"
223..1806
/gene="VC2246"
223..1806
/gene="VC2247"
223..1806
/gene="VC2248"
223..1806
/gene="VC2249"
223..1806
/gene="VC2250"
223..1806
/gene="VC2251"
223..1806
/gene="VC2252"
223..1806
/gene="VC2253"
223..1806
/gene="VC2254"
223..1806
/gene="VC2255"
223..1806
/gene="VC2256"
223..1806
/gene="VC2257"
223..1806
/gene="VC2258"
223..1806
/gene="VC2259"
223..1806
/gene="VC2260"
223..1806
/gene="VC2261"
223..1806
/gene="VC2262"
223..1806
/gene="VC2263"
223..1806
/gene="VC2264"
223..1806
/gene="VC2265"
223..1806
/gene="VC2266"
223..1806
/gene="VC2267"
223..1806
/gene="VC2268"
223..1806
/gene="VC2269"
223..1806
/gene="VC2270"
223..1806
/gene="VC2271"
223..1806
/gene="VC2272"
223..1806
/gene="VC2273"
223..1806
/gene="VC2274"
223..1806
/gene="VC2275"
223..1806
/gene="VC2276"
223..1806
/gene="VC2277"
223..1806
/gene="VC2278"
223..1806
/gene="VC2279"
223..1806
/gene="VC2280"
223..1806
/gene="VC2281"
223..1806
/gene="VC2282"
223..1806
/gene="VC2283"
223..1806
/gene="VC2284"
223..1806
/gene="VC2285"
223..1806
/gene="VC2286"
223..1806
/gene="VC2287"
223..1806
/gene="VC2288"
223..1806
/gene="VC2289"
223..1806
/gene="VC2290"
223..1806
/gene="VC2291"
223..1806
/gene="VC2292"
223..1806
/gene="VC2293"
223..1806
/gene="VC2294"
223..1806
/gene="VC2295"
223..1806
/gene="VC2296"
223..1806
/gene="VC2297"
223..1806
/gene="VC2298"
223..1806
/gene="VC2299"
223..1806
/gene="VC2300"
223..1806
/gene="VC2301"
223..1806
/gene="VC2302"
223..1806
/gene="VC2303"
223..1806
/gene="VC2304"
223..1806
/gene="VC2305"
223..1806
/gene="VC2306"
223..1806
/gene="VC2307"
223..1806
/gene="VC2308"
223..1806
/gene="VC2309"
223..1806
/gene="VC2310"
223..1806
/gene="VC2311"
223..1806
/gene="VC2312"
223..1806
/gene="VC2313"
223..1806
/gene="VC2314"
223..1806
/gene="VC2315"
223..1806
/gene="VC2316"
223..1806
/gene="VC2317"
223..1806
/gene="VC2318"
223..1806
/gene="VC2319"
223..1806
/gene="VC2320"
223..1806
/gene="VC2321"
223..1806
/gene="VC2322"
223..1806
/gene="VC2323"
223..1806
/gene="VC2324"
223..1806
/gene="VC2325"
223..1806
/gene="VC2326"
223..1806
/gene="VC2327"
223..1806
/gene="VC2328"
223..1806
/gene="VC2329"
223..1806
/gene="VC2330"
223..1806
/gene="VC2331"
223..1806
/gene="VC2332"
223..1806
/gene="VC2333"
223..1806
/gene="VC2334"
223..1806
/gene="VC2335"
223..1806
/gene="VC2336"
223..1806
/gene="VC2337"
223..1806
/gene="VC2338"
223..1806
/gene="VC2339"
223..1806
/gene="VC2340"
223..1806
/gene="VC2341"
223..1806
/gene="VC2342"
223..1806
/gene="VC2343"
223..1806
/gene="VC2344"
223..1806
/gene="VC2345"
223..1806
/gene="VC2346"
223..1806
/gene="VC2347"
223..1806
/gene="VC2348"
223..1806
/gene="VC2349"
223..1806
/gene="VC2350"
223..1806
/gene="VC2351"
223..1806
/gene="VC2352"
223..1806
/gene="VC2353"
223..1806
/gene="VC2354"
223..1806
/gene="VC2355"
223..1806
/gene="VC2356"
223..1806
/gene="VC2357"
223..1806
/gene="VC2358"
223..1806
/gene="VC2359"
223..1806
/gene="VC2360"
223..1806
/gene="VC2361"
223..1806
/gene="VC2362"
223..1806
/gene="VC2363"
223..1806
/gene="VC2364"
223..1806
/gene="VC2365"
223..1806
/gene="VC2366"
223..1806
/gene="VC2367"
223..1806
/gene="VC2368"
223..1806
/gene="VC2369"
223..1806
/gene="VC2370"
223..1806
/gene="VC2371"
223..1806
/gene="VC2372"
223..1806
/gene="VC2373"
223..1806
/gene="VC2374"
223..1806
/gene="VC2375"
223..1806
/gene="VC2376"
223..1806
/gene="VC2377"
223..1806
/gene="VC2378"
223..1806
/gene="VC2379"
223..1806
/gene="VC2380"
223..1806
/gene="VC2381"
223..1806
/gene="VC2382"
223..1806
/gene="VC2383"
223..1806
/gene="VC2384"
223..1806
/gene="VC2385"
223..1806
/gene="VC2386"
223..1806
/gene="VC2387"
223..1806
/gene="VC2388"
223..1806
/gene="VC2389"
223..1806
/gene="VC2390"
223..1806
/gene="VC2391"
223..1806
/gene="VC2392"
223..1806
/gene="VC2393"
223..1806
/gene="VC2394"
223..1806
/gene="VC2395"
223..1806
/gene="VC2396"
223..1806
/gene="VC2397"
223..1806
/gene="VC2398"
223..1806
/gene="VC2399"
223..1806
/gene="VC2400"
223..1806
/gene="VC2401"
223..1806
/gene="VC2402"
223..1806
/gene="VC2403"
223..1806
/gene="VC2404"
223..1806
/gene="VC2405"
223..1806
/gene="VC2406"
223..1806
/gene="VC2407"
223..1806
/gene="VC2408"
223..1806
/gene="VC2409"
223..1806
/gene="VC2410"
223..1806
/gene="VC2411"
223..1806
/gene="VC2412"
223..1806
/gene="VC2413"
223..1806
/gene="VC2414"
223..1806
/gene="VC2415"
223..1806
/gene="VC2416"
223..1806
/gene="VC2417"
223..1806
/gene="VC2418"
223..1806
/gene="VC2419"
223..1806
/gene="VC2420"
223..1806
/gene="VC2421"
223..1806
/gene="VC2422"
223..1806
/gene="VC2423"
223..1806
/gene="VC2424"
223..1806
/gene="VC2425"
223..1806
/gene="VC2426"
223..1806
/gene="VC2427"
223..1806
/gene="VC2428"
223..1806
/gene="VC2429"
223..1806
/gene="VC2430"
223..1806
/gene="VC2431"
223..1806
/gene="VC2432"
223..1806
/gene="VC2433"
223..1806
/gene="VC2434"
223..1806
/gene="VC2435"
223..1806
/gene="VC2436"
223..1806
/gene="VC2437"
223..1806
/gene="VC2438"
223..1806
/gene="VC2439"
223..1806
/gene="VC2440"
223..1806
/gene="VC2441"
223..1806
/gene="VC2442"
223..1806
/gene="VC2443"
223..1806
/gene="VC2444"
223..1806
/gene="VC2445"
223..1806
/gene="VC2446"
223..1806
/gene="VC2447"
223..1806
/gene="VC2448"
223..1806
/gene="VC2449"
223..1806
/gene="VC2450"
223..1806
/gene="VC2451"
223..1806
/gene="VC2452"
223..1806
/gene="VC2453"
223..1806
/gene="VC2454"
223..1806
/gene="VC2455"
223..1806
/gene="VC2456"
223..1806
/gene="VC2457"
223..1806
/gene="VC2458"
223..1806
/gene="VC2459"
223..1806
/gene="VC2460"
223..1806
/gene="VC2461"
223..1806
/gene="VC2462"
223..1806
/gene="VC2463"
223..1806
/gene="VC2464"
223..1806
/gene="VC2465"
223..1806
/gene="VC2466"
223..1806
/gene="VC2467"
223..1806
/gene="VC2468"
223..1806
/gene="VC2469"
223..1806
/gene="VC2470"
223..1806
/gene="VC2471"
223..1806
/gene="VC2472"
223..1806
/gene="VC2473"
223..1806
/gene="VC2474"
223..1806
/gene="VC2475"
223..1806
/gene="VC2476"
223..1806
/gene="VC2477"
223..1806
/gene="VC2478"
223..1806
/gene="VC2479"
223..1806
/gene="VC2480"
223..1806
/gene="VC2481"
223..1806
/gene="VC2482"
223..1806
/gene="VC2483"
223..1806
/gene="VC2484"
223..1806
/gene="VC2485"
223..1806
/gene="VC2486"
223..1806
/gene="VC2487"
223..1806
/gene="VC2488"
223..1806
/gene="VC2489"
223..1806
/gene="VC2490"
223..1806
/gene="VC2491"
223..1806
/gene="VC2492"
223..1806
/gene="VC2493"
223..1806
/gene="VC2494"
223..1806
/gene="VC2495"
223..1806
/gene="VC2496"
223..1806
/gene="VC2497"
223..1806
/gene="VC2498"
223..1806
/gene="VC2499"
223..1806
/gene="VC2500"
223..1806
/gene="VC2501"
223..1806
/gene="VC2502"
223..1806
/gene="VC2503"
223..1806
/gene="VC2504"
223..1806
/gene="VC2505"
223..1806
/gene="VC2506"
223..1806
/gene="VC2507"
223..1806
/gene="VC2508"
223..1806
/gene="VC2509"
223..1806
/gene="VC2510"
223..1806
/gene="VC2511"
223..1806
/gene="VC2512"
223..1806
/gene="VC2513"
223..1806
/gene="VC2514"
223..1806
/gene="VC2515"
223..1806
/gene="VC2516"
223..1806
/gene="VC2517"
223..1806
/gene="VC2518"
223..1806
/gene="VC2519"
223..1806
/gene="VC2520"
223..1806
/gene="VC2521"
223..1806
/gene="VC2522"
223..1806
/gene="VC2523"
223..1806
/gene="VC2524"
223..1806
/gene="VC2525"
223..1806
/gene="VC2526"
223..1806
/gene="VC2527"
223..1806
/gene="VC2528"
223..1806
/gene="VC2529"
223..1806
/gene="VC2530"
223..1806
/gene="VC2531"
223..1806
/gene="VC2532"
223..1806
/gene="VC2533"
223..1806
/gene="VC2534"
223..1806
/gene="VC2535"
223..1806
/gene="VC2536"
223..1806
/gene="VC2537"
223..1806
/gene="VC2538"
223..1806
/gene="VC2539"
223..1806
/gene="VC2540"
223..1806
/gene="VC2541"
223..1806
/gene="VC2542"
223..1806
/gene="VC2543"
223..1806
/gene="VC2544"
223..1806
/gene="VC2545"
223..1806
/gene="VC2546"
223..1806
/gene="VC2547"
223..1806
/gene="VC2548"
223..1806
/gene="VC2549"
223..1806
/gene="VC2550"
223..1806
/gene="VC2551"
223..1806
/gene="VC2552"
223..1806
/gene="VC2553"
223..1806
/gene="VC2554"
223..1806
/gene="VC2555"
223..1806
/gene="VC2556"
223..1806
/gene="VC2557"
223..1806
/gene="VC2558"
223..1806
/gene="VC2559"
223..1806
/gene="VC2560"
223..1806
/gene="VC2561"
223..1806
/gene="VC2562"
223..1806
/gene="VC2563"
223..1806
/gene="VC2564"
223..1806
/gene="VC2565"
223..1806
/gene="VC2566"
223..1806
/gene="VC2567"
223..1806
/gene="VC2568"
223..1806
/gene="VC2569"
223..1806
/gene="VC2570"
223..1806
/gene="VC2571"
223..1806
/gene="VC2572"
223..1806
/gene="VC2573"
223..1806
/gene="VC2574"
223..1806
/gene="VC2575"
223..1806
/gene="VC2576"
223..1806
/gene="VC2577"
223..1806
/gene="VC2578"
223..1806
/gene="VC2579"
223..1806
/gene="VC2580"
223..1806
/gene="VC2581"
223..1806
/gene="VC2582"
223..1806
/gene="VC2583"
223..1806
/gene="VC2584"
223..1806
/gene="VC2585"
223..1806
/gene="VC2586"
223..1806
/gene="VC2587"
223..1806
/gene="VC2588"
223..1806
/gene="VC2589"
223..1806
/gene="VC2590"
223..1806
/gene="VC2591"
223..1806
/gene="VC2592"
223..1806
/gene="VC2593"
223..1806
/gene="VC2594"
223..1806
/g
```

```
/gene="VC2016"
/notes="similar to GB:L42023 SP:P44719 PID:1003796
PID:1222388 FID:1204706; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="thymidylate kinase"
/protein_id="AA095164.1"
/db_xref="GI:9656560"
/translation="MRKRTIVIELEGAGKSTAIQVVETLQONGIDHITRERPGGT
LLAKRULVKEHPGEEQLDITTELLSVTAARVQLVENVIKPALARGEVWVGDRHDS
SQYGGGRIQSTQKQALGDFKPLDLYLIDPKLGLERARGELDRTERKM
DISFFRAREYRLANSDDSVVIDAQSISQVTADIRRALQDWLSQVNRV"
/complement(6253..7269)
/gene="VC2017"
/complement(6253..7269)
/gene="VC2017"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA095165.1"
/db_xref="GI:9656561"
/translation="MKKXULVILVIGIVAGSYFYVVKQMDQYLAQPLMIDEAQIVT
IASGTLRELALQTEQAKQDSFVAEVRHHPKSKTKAGTYKQDPMSLEQALAL
LVSGKHQSITVEGSRFOEMRDILASNNETIQTLTGLTIDIAKAGIEHEKLEGL
FLAETHYFTKGTSDVRLKRAKQLEKFLQVMEHQEGLPIQTYEALILASIIEXE
TSAEBERERIAAVFINRLKRMLEQDPTVIYMGSETYDGNIRKKDLARTPYNYVI
NGLPPTPIAMPGEASIVAAALNPQSDYLVFVAGSGEGHNFSTLADHNRVAYLKL
RTKQ"
/complement(7266..8099)
/gene="VC2018"
/complement(7266..8099)
/gene="VC2018"
/notes="similar to PID:1173841 PID:1173844; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="4-aminio-4-deoxychorismate lyase"
/protein_id="AA095166.1"
/db_xref="GI:9656562"
/translation="MSRFTLWSESMYVNGQRNRPVPHDRSFQYGDGFTTLTKEG
QVQMSHSHKQAQCLDIHLPEPNWDRVQGLQSMILPOEKAGLKHISRLGLGRGY
SPTQVSESIVTISAPFAHPYQAWRDKLAVGICCOORGLNPLLAGHKHNNLEQILL
KXEMDNAGDDGVCLDINKVLETTAANTFWCRDGTMTFPCRHRAGVAVRRQLEL
AAQQEPIVIDEFTLELLSAREEVFIINALLEVAPEVTQIQQRQLTIGSMTRRFQESSN
S"
/complement(8147..9391)
/gene="VC2019"
/complement(8147..9391)
/gene="VC2019"
/notes="similar to PID:1173841 SP:P55338 PID:1173843;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="3-oxoacyl-(acyl-carrier-protein) synthase II"
/protein_id="AA095167.1"
/db_xref="GI:9656563"
/translation="MSKRRVVVTGMCLSPVCNTVSESSKALLAQSGIVNIEHPDIT
NSTRPAGLVKGFDEQNSKKDKARDLFIQIGIAAGLQALDESLGVNEENRARIQ
VAIGSGLELIETHQALIEKPRKVPFFVPTSTVIMNAGNLSIMRGLRGNIAI
STACTTGLNIIHAARMIAYGADAWAGGAEKASTPLGMAGFGAAKALSTRNDEPQK
ASRPMDKRDGFLVGAGTIGVLEVEYHAKARGAKIYAEVVGFMGSDAYHMTSPED
GGGALAMEAARNADVGTQIGYVNAHGTSTPAGDVAEVKGIKRALGEAGTKOVLVS
STKSMTGHLILGAAGSVEAIIITWMSLVQDVQVPTINLDNPEELGVDLVPHVARKVESM
EYAMNCSFGFGTNGSLIFKRM"
/complement(9480..9812)
/gene="VC2020"
/complement(9480..9812)
/gene="VC2020"
/notes="similar to PID:1173841 SP:P55337 PID:1173842;
identified by sequence similarity; putative"
/codon_start=1
```

```
/transl_table=11
/product="acyl carrier protein"
/protein_id="AA095168.1"
/db_xref="GI:9656564"
/translation="MVRPETSRLQSLVRINYGIALGEICKGKEMNIEBVRVKII
VEQLGVDBAEVKNESSFVEDLGADSLDTVELWALBEEFDETEIPDEAEKITTQAAI
DYVTSNAQ"
/complement(9952..10698)
/gene="VC2021"
/complement(9952..10698)
/gene="VC2021"
Query Match 23.8%; Score 219.6; DB 1; Length 14530;
Best Local Similarity 54.5%; Pred. No. 3.9e-49;
Matches 507; Conservative 0; Mismatches 414; Indels 9; Gaps 3;

QY 1 ATGACTAAACAGCGCTTTTATTTGCTGGTCAAGTGCCCAAGTATCTAGGATCGGACGG 60
DB 11637 ATGAGTAAGTTTGTATCGTATTTCCAGGTCAAGGCTCGAAGCAGTAGGTATCTGGCT 11578
QY 61 GATTTCATGATCAGTATCGGATTTCAAGAAACAGATTGATCGAGCGAGTCAGGTGCTC 120
DB 11577 GACCTTGGCGAGGAGTATGCTGTGTTAAACAACATTCGCCGAAGCTTCAGAAGTGCCT 11518
QY 121 GGTATGATTTACGTTATCTCATC---GATACGGAAGAACACAACTCAATCAGACCGGC 177
DB 11517 GGTTACGATCTGTGGCGCTGGTTCAAGATGGCCCTGTGGAAGATCTCAACCAAACTTC 11458
QY 178 TATACGAAACAGCAGCATCTAGCGACTTCGGTGTGCTATCTACCGTTTATTCAGAAAG 237
DB 11457 CGTACTCAACTGCTGTGCTTGGCGCTCTGTTCGGAATTTGGCGGTATGCGCAGCACTG 11398
QY 238 GG---CTATCAGCCTGATATGTTGCTGTGTTGTTCTCTTGGAGAAATCTCTGCCCTTGGTG 294
DB 11397 GGTCTTGAGCAACCTCGCGTTTATAGTGGTCACAGCTGGGTGAATATTGAGCACTGGTA 11338
QY 295 GCAAGCGCGCTTGGATTTTGAAGATGCGGTTCCTTGGTAGTAGTACGCTGAGCGCTAT 354
DB 11337 TGTGCGGCGGTGATTGATTTTAAACAAGCGATCAAGCTGTTGAGCTGCGGTGTCATTTG 11278
QY 355 ATGAAAGAAAGCGGCTCTGCTGACTTGGCAAGATGTTAGCAGTCTCAATACGCCCACTA 414
DB 11277 ATGCAACAAGCGGTGCTCGCGGTACGGTGCATATGACCGATCATCGGTCTAGAAGT 11218
QY 415 GAGTCTATTGAAGAAGCCTGTCAAAAAGCTTCTGAACTTGGAGTGTGTTACTTCAGCACAAC 474
DB 11217 GAAAGCGATTGCTAAAGCGCTGTGCTGACGCGCGCAGGTGAAGTGGTTTCTCTGTAAAC 11158
QY 475 TATAACACCTGCACAAAATCGTCAATTTGCTGGAGAAGTGTGTCAGTTGATCGAGCGGTT 534
DB 11157 TTTAACTCACAGGCGCAAGTGGTTATTGCTGGTCAAAAAGATGCGGTGAGCGTGGCGGC 11098
QY 535 GAATTTTGAAGAAGCAGGTGCGCAAGCGTTGATTCCTCTTTAAGGTGTCAGGTCCCTTT 594
DB 11097 GTTCTGTGTTAAAGAAGCGGCGGCAACGTCGCGCTCTGCGCAGTTTCCGTACCATCA 11038
QY 595 CACACCGCTCTCCTTTGAGCGCTGTAGCCAGAAACTAGCTCAAACTCTAGCTCAGGTAAGT 654
DB 11037 CACTGCGCGTTGATGAAGCGCTGTGCGCGATGAATTTGCAAAAACCTCTAGCAGAGCTTGA 10978
QY 655 TTTTCAGATTTTACTTGTCCCTCTAGTCGCGCAATACAGAAGCT---GCTGTGATGCAAAA 711
DB 10977 TTCAATGCACCACAAAATCCGGTTCATCAATAACGTTGATTTGCTGGCTGAAACGGATCCG 10918
QY 712 GAGGACATCTCAGCTCTTTGACGCGTCAGTCAAGGAACCGCTTGGTTCTTATGAAGT 771
DB 10917 GTAAAAATTAAGATGCGTTGATTCGTCAACTCTATAGCCAGTTCGTTGGACTGAATGC 10858
QY 772 ATTGGGCTATGCAAGAAGCAGGCATAAGCAACTTTTTCAGAGATTGACCGGGAAGATC 831
DB 10857 GTTGCAACAATGAGCGCACAGGTGTCGAAAGCTGATTGAATGGGCGCGGTAAAGTA 10798
QY 832 TTGTCAGGTTTTGTTAAAAAANTGATCAAACTGCTCACTAGTCTATGTTGGAAGATCAA 891
```

Db 10797 TTGACTGGTCTACAAACGTTATGTAACACCTAGAGGTGCGCAGTCAATGACGTG 10738
 QY 892 GCGAGTTTAGTACACTTTTACAAAAATAG 921
 Db 10737 GCTTCTTTGGATGCGGTGAATAACAACAG 10708

RESULT 14
 AEO13105/c
 LOCUS AEO13105/c
 DEFINITION Thermoanaerobacter tengcongensis strain MB4T, section 132 of 244 of the complete genome.
 ACCESSION AEO13105 AEO08691
 VERSION AEO13105.1 GI:20516482
 KEYWORDS
 SOURCE Thermoanaerobacter tengcongensis.
 ORGANISM Thermoanaerobacter tengcongensis
 Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 Thermoanaerobacteriaceae; Thermoanaerobacter.
 REFERENCE 1 (bases 1 to 11845)
 AUTHORS Bao, Q., Tian, Y., Li, W., Xu, Z., Xuan, Z., Hu, S., Dong, W., Yang, J., Chen, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y., Ling, L., Tan, H., Chen, R., Wang, J., Yu, J., and Yang, H.
 A Complete Sequence of the T. tengcongensis Genome
 Genome Res. 12 (5), 689-700 (2002)
 MEDLINE 21992816
 PUBMED 11997336
 REFERENCE 2 (bases 1 to 11845)
 AUTHORS Bao, Q., Xu, Z., Hu, S., Dong, W., Chen, Y., Wang, J., Yu, J., and Yang, H.
 Direct Submission
 TITLE Submitted (07-OCT-2001) Beijing Genomics Institute/Genomics and Bioinformatics Center, Institute of Genetics and Development, Chinese Academy of Sciences, Beijing Airport Industrial Zone B8, Beijing 101300, China
 JOURNAL Beijing 101300, China
 REFERENCE 3 (bases 1 to 11845)
 AUTHORS Li, W., Xuan, Z., Yang, J., Ling, L. and Chen, R.
 Direct Submission
 TITLE Submitted (07-OCT-2001) Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China
 JOURNAL Sciences, Beijing 100101, China
 REFERENCE 4 (bases 1 to 11845)
 AUTHORS Tian, Y., Xue, Y., Xu, Y., Lai, X., Huang, L., Dong, X., Ma, Y. and Tan, H.
 Direct Submission
 TITLE Submitted (07-OCT-2001) Institute of Microbiology, Chinese Academy of Sciences, Beijing 100080, China
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. 11845
 /organism="Thermoanaerobacter tengcongensis"
 /strain="MB4T"
 /db_xref="taxon:119072"
 complement(76. 1153)
 /gene="ELP3"
 /note="TIE1468"
 complement(76. 1140)
 /note="ELP3"
 /gene="ELP3"
 /note="Best Blastp hit = gi|7462101|pir||A72295 conserved hypothetical protein - Thermotoga maritima (strain MSB8) gi|4981648|gb|AAD36179.1|AEO01769.9 (AEO01769) conserved hypothetical protein [Thermotoga maritima], score 202, E-value 6.00E-51"
 /codon_start=1
 /product="3-oxoacyl- (acyl-carrier-protein) synthase"
 /translation="MNRVVVTGLGVISPLGNTVERFNSLNGENGIGLVTKFDTSEF PTKVAAEVKDFDFTLYIDKKEARRMDRFIQFAVASTKLAEDAGIELDKVDLNRFGVI YGTGGTTEFENQKVMYKGFSPFFIPMIANMAAGQIAITFGLKGINETIVN ACATSTNAIGAEFAKAIQRGDADLIVTGTAAITVNSPAGFCAMKAMTNPDPSTASR PDLNADCFVMGEGSATLILSLHAQKRGAKIYAEIIGYGATDDAHITAPAGEG ARAAEALDKGVSYDMVDVINAHGTSTETENDKFTMAIKVFKDHAYKLCVSSNKS MIGHLLGAGAVEAVATVDTINAHGTSTETENDKFTMAIKVFKDHAYKLCVSSNKS NSFGFGHNATLLFFKFF"
 complement(1855. 2325)
 /gene="FabB"
 /note="Pfam match to entry ketoacyl-synt C, Beta-ketoacyl synthase, C-terminal domain, score 238.8, E-value 7.70E-68"
 complement(2347. 3078)
 /gene="FabB"
 /note="Pfam match to entry ketoacyl-synt, Beta-ketoacyl synthase, N-terminal domain, score 295.3, E-value 7.70E-85"
 complement(3088. 3092)
 /gene="FabB"
 complement(3111. 3127)
 terminator

terminator
 /note="putative rho-independent transcription terminator"
 complement(1137. 1837)
 /gene="Rnc"
 /note="TIE1469"
 complement(1137. 1823)
 /gene="Rnc"
 /note="Best Blastp hit = gi|10175109|dbj|BAB06208.1| (AEO01515) ribonuclease III [Bacillus halodurans], score 193, E-value 1.00E-48"
 /codon_start=1
 /transl_table=11
 /product="dsRNA-specific ribonuclease"
 /protein_id="AAM24691.1"
 /db_xref="GI:20516484"
 /translation="MLTYLEQKINYEFDKDTLLLEALTHSSWAHEGKNEKVNLERLEF LGDSVLISVISYLYKNRKDEGLSKYRAIVCEPSLARCARKELQSGFPMRGKGE EISGGRDRDSILADAMEALLAAVYLDGGLAEVRRVLDLFEIIDEVLKGIYRDYKT RIQEVVQSMNEVGKITYELVEBIGPDKNTFTVQVIGDVVLGIGGQKSKSEEQAAAM EALSUKLGLIK"
 complement(1149. 1153)
 /gene="ELP3"
 complement(1152. 1352)
 /gene="Rnc"
 /note="Pfam match to entry dsrm, Double-stranded RNA binding motif, score 85.2, E-value 1.30E-21"
 complement(1434. 1706)
 /gene="Rnc"
 /note="Pfam match to entry Ribonuclease_3, RNase3 domain, score 139, E-value 8.40E-38"
 complement(1833. 1837)
 /gene="Rnc"
 complement(1849. 3092)
 /gene="FabB"
 /note="TIE1470"
 complement(1849. 3078)
 /gene="FabB"
 /note="Best Blastp hit = gi|7433750|pir||G69842 3-oxoacyl- (acyl-carrier-protein) synthase (EC 2.3.1.41) - Bacillus subtilis gi|2633471|emb|CAB12975.1| (Z99109) similar to 3-oxoacyl- acyl-carrier protein synthase [Bacillus subtilis] gi|2633488|emb|CAB12991.1| (Z99110) similar to 3-oxoacyl- acyl-carrier protein synthase [Bacillus subtilis], score 496, E-value 1.00E-139"
 /codon_start=1
 /transl_table=11
 /product="3-oxoacyl- (acyl-carrier-protein) synthase"
 /protein_id="AAM24692.1"
 /db_xref="GI:20516485"
 /translation="MNRVVVTGLGVISPLGNTVERFNSLNGENGIGLVTKFDTSEF PTKVAAEVKDFDFTLYIDKKEARRMDRFIQFAVASTKLAEDAGIELDKVDLNRFGVI YGTGGTTEFENQKVMYKGFSPFFIPMIANMAAGQIAITFGLKGINETIVN ACATSTNAIGAEFAKAIQRGDADLIVTGTAAITVNSPAGFCAMKAMTNPDPSTASR PDLNADCFVMGEGSATLILSLHAQKRGAKIYAEIIGYGATDDAHITAPAGEG ARAAEALDKGVSYDMVDVINAHGTSTETENDKFTMAIKVFKDHAYKLCVSSNKS MIGHLLGAGAVEAVATVDTINAHGTSTETENDKFTMAIKVFKDHAYKLCVSSNKS NSFGFGHNATLLFFKFF"
 complement(1855. 2325)
 /gene="FabB"
 /note="Pfam match to entry ketoacyl-synt C, Beta-ketoacyl synthase, C-terminal domain, score 238.8, E-value 7.70E-68"
 complement(2347. 3078)
 /gene="FabB"
 /note="Pfam match to entry ketoacyl-synt, Beta-ketoacyl synthase, N-terminal domain, score 295.3, E-value 7.70E-85"
 complement(3088. 3092)
 /gene="FabB"
 complement(3111. 3127)
 terminator

```

/note="putative rho-independent transcription terminator"
complement(3133..3432)
/gene="Acpp"
/note="TTE1471"
complement(3133..3432)
/gene="Acpp"
/note="Best Blastp hit = gi|3912967|sp|O67611|ACP_AQUAE
ACYL CARRIER PROTEIN (ACP) gi|7442262|pr||A70448_acyl
carrier protein Aquifex aeolicus
gi|2984024|gb|AAC07567.1| (AE000752) acyl carrier protein
[Aquifex aeolicus], score 75.5, E-value 2.00E-13"
/codon_start=1
/transl_table=11
/product="Acyl carrier protein"
/protein_id="AAW24693.1"
/db_xref="GI:20516486"
/translation="MLTEEWLNCNIIISVQLRGRGKMFVKRVNIIARQLGIDPREI
TMSSFFIDLGASLDIVELIMALEEFIDPEDAEKIKTVGDVVEYLSNVE"
complement(3151..3354)
/gene="Acpp"
/note="Pfam match to entry pp-binding, Phosphopantetheine
attachment site, score 95.2, E-value 9.80E-25"
complement(3407..4163)
/gene="FabG3"
/note="TTE1472"
complement(3407..4150)
/gene="FabG3"
/note="Best Blastp hit = gi|1502421|gb|AAC44307.1|
(U59433) 3-ketoacyl-acyl carrier protein reductase
[Bacillus subtilis], score 265, E-value 6.00E-70"
/codon_start=1
/transl_table=11
/product="Dehydrogenases with different specificities
(related to short-chain alcohol dehydrogenases)"
/protein_id="AAW24694.1"
/db_xref="GI:20516487"
/translation="WDEKRVAVTGSRGIGRAIARLAKDGFNIATYVKKDKSAE
EVEBKKGGDALAKDSKVEHEKAVEKVEIEFGSDVIVNNAGITKONLILKM
EEEDWDVIVNLKGFANVIFASKYMKRKGKIINISVWGLMGNVQOANYASKA
GIILTKSAKEIASGITVNAVPGFIETDMTNVLKEDIKEAMLSKSIPLKRGKPEE
VAEYVFLASASDYITGOVINVDGGMVM"
complement(3428..3520)
/gene="FabG3"
/note="Pfam match to entry adh short_C2, Short chain
dehydrogenase/reductase C-terminus, score 59.2, E-value
3.40E-15"
complement(3578..4135)
/gene="FabG3"
/note="Pfam match to entry adh short, short chain
dehydrogenase, score 283.1, E-value 3.60E-81"
complement(4154..5093)
/gene="FabD"
/note="TTE1473"
complement(4154..5083)
/gene="FabD"
/note="Best Blastp hit = gi|7404378|sp|P71019|FABD_BACSU
MALONYL COA-ACYL CARRIER PROTEIN TRANSACYLASE (MCT)-
gi|743373|pr||H6920 [acyl-carrier-protein]
S-malonyltransferase (EC 2.3.1.39) fabd [validated] -
Bacillus subtilis gi|2337819|emb|CAA74249.1| (Y113937)
putative FabD protein [Bacillus subtilis]
gi|2633962|emb|CAB13463.1| (299112) malonyl CoA-acyl
carrier protein transacylase [Bacillus subtilis], score
276, E-value 2.00E-73"

```

Query Match 23.1%; Score 212.4; DB 1; Length 11845;
Best Local Similarity 53.0%; Pred. No. 3.7e-47;
Matches 478; Conservative 0; Mismatches 421; Indels 3; Gaps 1;

Qy 1 ATGACTAAACAGCCTTTTATTGCTGGTCAAGTGCCCGAGTATCTAGGATGGGACGG 60
|||||
Db 5086 AGGATGAATAGCTTTTATATACCCGGCCAGGTGCCCACTACCGAGGATGGGAAG 5027

```

Qy 61 GATTTCTATGATCATGATCCCGATTCTCAAGAAACGATTGATCGAGCGAGTCAGTGCTC 120
|||
Db 5026 GAAATATACGAAAAATACGAAGAACGAAAGAAATTTTGAAGAGCAGATGAGGCTTTA 4967
|||
Qy 121 GGTATGATTTTACGTTATCT---CATCGATACGGAAAGACAAACTCAATCAGACCCGC 177
|||
Db 4966 GGTTTTAAATATCTAACTGTGCTTTGAAGGGCCCTGAAGAGAGCTTATGAAGACGGAG 4907
|||
Qy 178 TATAGCAACAGCCANTCTTAGCGACTTCGGTTGCTATCTACCGTTTATTTGCAAGAAAG 237
|||
Db 4906 AACACACAGCCGCCCATTTTGACAGTTTCAGTGGCTTTCAGCAAGGGTTTTCAGAAAAA 4847
|||
Qy 238 GCCTATCAGCCGATATGTTGCTGTTGCTCTTTCGAGAAATATCTCTGCTTGGTGCA 297
|||
Db 4846 GGTGTAARAACTGATGTGACAGCAGTTTAAAGTTAGAGAGTATTCTCTCTTGTACTG 4787
|||
Qy 298 AGCGCGCTTGATTTTGAAGATCGGTTGCCCTTGGTAGTAAAGCTGGAGCCCTATATG 357
|||
Db 4786 GCAGAAAGCTTTAGATTTTGAAGATCAGTGAGACTTGTCAAAAAGAGAGAGGAAATACATG 4727
|||
Qy 358 GAAGAAAGCGGCTCCTGCTGACTCTGCAAGATGGTAGCAGTTCTCAATACGCCACTAGAG 417
|||
Db 4726 CAGGAAGTGGTCCCAAGAGAGTCCGCCAATATGGCAGCAATACITGGACTTCCCAATGAG 4667
|||
Qy 418 GTCATTGAAGAGCCCTGTCAAAAAGCTTCTGAACCTTGGAGTGGTTTACTCCAGCCCACTAT 477
|||
Db 4666 GAAAGTTGAGGAAATATGCCGATTCATCAGAGGTCGGTGTGGTAGAACTTCTTAACATAC 4607
|||
Qy 478 AACACACCTGCAACAATCGTCTGCTGAGAGTGGTTGCGAGTGTGATCGAGCGGTTGAA 537
|||
Db 4606 AACTCTCCGGTCAATTAGTGGTCTCTGGGAAGTTAAGGCTGTAGAAAGAGCTGTGAA 4547
|||
Qy 538 CTTTTGCAAGAGCAGGTGCCAAACGCTTGATTCCTCTTAAGGTGTCAAGTCCCTTTCCAC 597
|||
Db 4546 CTAGCGAAGGAGAGAGGAGCTTAAGAAACAGTAGTGTTCGAGTAAGTGCCTTTCCAC 4487
|||
Qy 598 ACCGCTCTCTTGAGCCTCTAGCCAGAACTAGCTGAACTAGCTCAGCTCAGSTAAGTTT 657
|||
Db 4486 TGCAGCATGTTGAAAGAGAGCTGGAGAACTTCTAGCAAAAGGAATTTGGATAAAGTAGAGATA 4427
|||
Qy 658 TCAGATTTTACTTGTGCCCTAGTCGGCAATACAGAAGCTGCTGTGATGCAAAAAGAGGAC 717
|||
Db 4426 AAGGACTTGAAGTTCTCTGTATATCAATGTTACTGCCGATTACGTCGCAAAAAGATAAG 4367
|||
Qy 718 ATTGCTAGCTCTTGACCGCTCAGCTCAAGGAAACCGGTTCTGTTCTATGAAAGTATTGGG 777
|||
Db 4366 GTTAAAGAAATTACTCATAAACAGGTGAGTCATCCCGTCTCTGGAGCAGTCTGTGAGA 4307
|||
Qy 778 GTCATGCAAGAGCAGGCGATTAAGCAACTTATCGAGATTGACCGGGGAAAGTCTTGTCAC 837
|||
Db 4306 AAATGATAGAAGACGGTGTGGATCTTTCTAGAAATAGGACCTGGAAAGACTTTGAGC 4247
|||
Qy 838 GGTTTTGTAAAAAAATTTGATCAAACTGCTCAGCTAGCTCATGTGGAAGATCAAGCGAGT 897
|||
Db 4246 GGATTTGTCAAGAAAAATTTGATAGAAGCAGGACTGTTTGAATTTTGGAGATGAGGAGTCT 4187
|||
Qy 898 TT 899
Db 4186 CT 4185

```

RESULT 15
AX432728
LOCUS AX432728
DEFINITION Sequence 1143 from Patent WO02291113.
ACCESSION AX432728
VERSION AX432728.1 GI:21657532
KEYWORDS
SOURCE Bacillus licheniformis.
ORGANISM Bacillus licheniformis.
REFERENCE 1 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
AUTHORS Berka, R. and Clausen, I. G.

TITLE Methods for monitoring multiple gene expression

JOURNAL Patent: WO 0229113-A1143 11-APR-2002;

Novozymes Biotech, Inc. [US]; Novozymes A/S (DK)

FEATURES Location/Qualifiers

source

1. .950

/organism="Bacillus licheniformis"

/db_xref="taxon:1402"

BASE COUNT 276 a 223 c 259 g 192 t

ORIGIN

Query Match

Best Local Similarity 22.1%; Score 204; DB 6; Length 950;

Matches 474; Conservative 0; Mismatches 385; Indels 9; Gaps 3;

QY 1 ATGACTAAACAGCCCTTTTATTGTGTCAGGTGCCAGTATCTAGGGATGGAGCG 60

Db 1 ATGGCAAGATTGCTTTTCTATTCCGGGCCAAGGTTCGACGATATCGCATGGACAC 60

QY 61 GATTTCATGATCAGTATCGGATGTCAAGAAACGATTGTCGAGCGAGTCAGGTGCTC 120

Db 61 GAATTGTATGAAAAGAACCGGATGCGAAGAGATTTTGAAGAAGCGGATCAAAACGCTT 120

QY 121 GGTATGATTACGTTTATCTCA---TCGATACGGAAGACAAACTCAATCAGACCCGC 177

Db 121 GAACAAACTGACACCTCATGTTGAAGGGATGCAAGAACTGACGCTTACATAC 180

QY 178 TATACGCAACACCCATTCTAGGACTTCGGTTGTATCTACCGTTTATTCGAAGAAAG 237

Db 181 AACGCGCAGCAAGCCCTTTTAACGGCGAGCATCGCAGCCCTTGAAGAACTGAAGGAATAC 240

QY 238 GGTATCAGCTGATATGTTGCTGTTGTTGTCTCTGGAGATATCTCTGCTTGGTGCA 297

Db 241 GGCATTAAAGCCGACTATCGGCGAGTACAGCTCGGCGAATACAGCGCATTTGGTCGT 300

QY 298 AGCGGCGCTTGATTTGAAGATGCGTTGCTTGGTAGCTAAGCTGGAGCCCTATATG 357

Db 301 GCGGCGCTTGTGCTTTAAAGATGCGTTTATGCGTCAGAAAGCGCGGGAATTCATG 360

QY 358 GAAGAGCGCTCTGCTGACTCTGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAG 417

Db 361 ATGAAGCGCTGCGCGCGGAGAGCGCGCATTCGCGCATGGAGCCAG 420

QY 418 GTCATTGAAGAGCCCTGTCAAAAGCTTCTGAACCTTGA---GTGGTTACTCCAGCCAAC 474

Db 421 GCGCTGAAGAGAGTGAAGGACGACAAATTTCCGAAGAGGAACCTTTGTCAGCTGCCAAT 480

QY 475 TATAACACAGCTGCACAAATGTCATTTGCTGGAGAGTGGTTGCAAGTTGATCGAGCGGTT 534

Db 481 TTGAATGCGCTGGGCAAAATCGTCATCTCGGAACAGCTTAAAGGCGTGGAGCTCGCTTCA 540

QY 535 GAACTTTTGAAGAGCAGGTGCAAAAGCTTGTATTCCTTTAAGGTGTCAGGTCCCTTT 594

Db 541 GAGCTTGAAGAGAGAGGCGCAAAAGCGCGATTCCTCTGAAAGTCAAGCGGCGGCTTC 600

QY 595 CACACCGCTCTCTTGAAGCTGCTAGCCAGAACTAGCTGAACCTCTAGCTCAGGTAAGT 654

Db 601 CATCTGAGCTGATGAAGCGGCGAGCTGATAAGCTTCGTGAAGTCTTGAATCGGTGACG 660

QY 655 TTTTCAGATTTTACTTGTCCCTAGTCGCAAA---TACAGAGCTGCTGTATGCAAAA 711

Db 661 ATCAAGAGCGAGCCATTCGGTGTCTCCAAAGTAAAGCGCGACTTTGTACGGATPAA 720

QY 712 GAGGACATTGCTAGCTTTGAGCGGTGAGGTCAAGAAACCGTTTCGTTTCTATGAAAGT 771

Db 721 GACGACATTGAAGATAAAGTGAATGAACAGCTGTATTCCTCTGTACGCTTTGAAGAAACA 780

QY 772 ATTGGGGTCATCAGAGAGCGGATAGCACTTTATCGAGATTGACCGGGAAGTC 831

Db 781 ATCAGCGGCTGATTGACGAAGCGGTACGACCTTCATTGAATCGGTCCCGGAAAGGTT 840

QY 832 TTGTCAGGTTTGTAAAAAAATTGATC 859

Db 841 TTGTCAGGCTTGTGAAGAAAGTGAACC 868

Search completed: June 11, 2003, 18:44:52
Job time : 2441 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2003, 17:03:56 ; Search time 264 Seconds
(without alignments)
7856.403 Million cell updates/sec

Title: US-09-308-397-1
Perfect score: 921
Sequence: 1 atgactaaacagcctttt.....tagcacttttagaaaaatag 921

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	921	100.0	921	19	AAV37239
2	913	99.1	921	22	AAH90730
3	913	99.1	19702	19	AAV52140
4	911.4	99.0	921	23	AAS55845
5	911.4	99.0	3580	19	AAV65243
6	908.4	98.6	945	22	AAH90843
7	529.4	57.5	547	19	AAV37241
8	523.2	56.8	1196	19	AAV38542
9	523.2	56.8	1199	18	AAV30801

C	10	523.2	56.8	1209	19	AAV37406
	11	439.4	47.7	942	24	ABN68641
	12	402	43.6	636	18	AAV30880
	13	388.8	42.2	2365589	24	ABA90521
	14	387	42.0	924	24	ABN68640
	15	387	42.0	2155561	24	ABN71527
	16	366	38.7	450	19	AAV37240
	17	344.4	37.4	942	23	AAS53186
	18	340.4	37.0	3656	20	AAV31514
	19	204	22.1	950	24	ABK73852
	20	204	22.1	960	24	ABK73795
	21	196.4	21.3	939	23	AAS53241
C	22	196.4	21.3	1830121	17	AAV42063
	23	185.6	20.2	930	23	AAS52391
	24	185.6	18.0	2301	24	ABQ70706
	25	162.4	17.6	319630	24	ABQ67194
	26	162.4	17.6	3011208	24	ABQ69244
C	27	160.4	17.4	2944528	24	ABA03041
	28	158	17.2	6021	20	AAV31395
	29	139.6	15.2	592	24	ABK78261
	30	136.6	14.8	4593	21	AAV81456
	31	136.6	14.8	34980	21	AAV21612
	32	136.6	14.8	837096	21	AAV81489
	33	130.4	14.2	936	24	AAV92524
	34	128	13.9	936	23	AAS54395
	35	127.6	13.9	906	23	AAS51823
	36	126.4	13.7	927	19	AAV31440
	37	126.4	13.7	999	22	AAV06204
	38	120.8	13.1	939	23	AAS54159
	39	118.4	12.9	1170	21	AAV51219
	40	116.8	12.7	672	22	AAV51330
	41	116.8	12.7	4155	22	AAV54909
C	42	112.6	12.2	954	24	ABQ90247
C	43	112	12.2	459	21	AAV21545
C	44	112	12.2	459	21	AAV81260
C	45	112	12.2	459	21	AAV33599

ALIGNMENTS

RESULT 1	AAV37239	AAV37239 standard; DNA; 921 BP.
ID	AAV37239	
XX	AAV37239;	
XX	AAV37239;	
DT	24-SEP-1998	(first entry)
DE	DNA encoding a FabD polypeptide.	
XX	FabD protein; malonyl-CoA:ACP family; diagnosis; infection; vaccine;	
KW	screen; ss.	
XX	Streptococcus pneumoniae.	
OS	Streptococcus pneumoniae.	
PH	Key	Location/Qualifiers
FT	CDS	1..921
FT		/tag= a
FT		/product= FabD
XX	WO9822133-A1.	
XX	28-MAY-1998.	
PD	14-NOV-1997;	97WO-US20992.
XX	18-NOV-1996;	96US-0031160.
XX	(SMK) SMITHKLINE BEECHAM CORP.	
PA	Gentry DR, Lonsdale JT, Payne DJ, Pearson SC, Van Aller G;	
XX		

Streptococcus pneu
Streptococcus poly
Streptococcus pneu
Genomic sequence o
Streptococcus poly
Streptococcus poly
N-terminal coding
Enterococcus faeca
Enterococcus faeca
Bacillus lichenifo
Bacillus lichenifo
Haemophilus influe
Haemophilus influe
E. coli DNA for ce
Listeria monocytog
Listeria innocua c
Listeria innocua D
Listeria monocytog
Enterococcus faeca
Bacillus clausii g
N meningitidis pa
Neisseria meningit
N meningitidis pa
Staphylococcus epi
Staphylococcus aur
Staphylococcus aur
Nucleotide sequenc
Pseudomonas aerugi
Staphylococcus aur
S. epidermidis ope
S. epidermidis gen
M. capsulatus gene
N meningitidis pa
N meningitidis pa
Neisseria meningit

DR WPI; 1998-312173/27.
XX P-PSDB; AAW60856.
PT New isolated Streptococcus pneumoniae FapD gene - used to develop
PT products for the diagnosis, prevention and treatment of bacterial
PT diseases, particularly S. pneumoniae infection
XX
XX
PS Claim 6; Pages 5-6; 45pp; English.
XX
CC The present sequence encodes a FapD protein of Streptococcus pneumoniae
CC 0100993 (NCIMB 40800). The novel FapD polypeptides are related to other
CC proteins of the malonyl-CoA:ACP family. The products can be used for the
CC diagnosis of Streptococcus pneumoniae infections. Vectors containing
CC the FapD DNA sequence can be administered directly to a mammal to
CC produce the FapD peptide to provoke an antibody/T-cell response in
CC order to prevent a disease. The peptide can be used to screen for
CC compounds which modulate its activity.
XX
SQ Sequence 921 BP; 247 A; 193 C; 230 G; 251 T; 0 other;
Query Match 100.0%; Score 921; DB 19; Length 921;
Best Local Similarity 100.0%; Pred. No. 3.1e-281;
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGACTAAACAGCCCTTTTATTTCTGCTCAAGGTGCCAGTATCTAGGATGGGACGG 60
Db |
QY 1 ATGACTAAACAGCCCTTTTATTTCTGCTCAAGGTGCCAGTATCTAGGATGGGACGG 60
QY 61 GATTCTATGATCAGTATCCGATTTGTCAGAAAGAGATGATCGAGCGATGAGTGTCTC 120
Db |
QY 61 GATTCTATGATCAGTATCCGATTTGTCAGAAAGAGATGATCGAGCGATGAGTGTCTC 120
QY 121 GGTATGATTTACGTTATCTATCATGATACGGAAGAGACAACTCAATCAGACCGGTAT 180
Db |
QY 181 ACGCAACAGCCATCTAGAGACTTCGGTGTCTATCTACCGTTTATTCGAAAGAAAGGGC 240
Db |
QY 181 ACGCAACAGCCATCTAGAGACTTCGGTGTCTATCTACCGTTTATTCGAAAGAAAGGGC 240
QY 241 TATCAGCTGATATGGTTGCTGGTTGCTCTTGGAGATATCTGCTGCTGGTGGCAAGC 300
Db |
QY 301 GCGCCTTTGGATTTTGAAGATCGGTTGCTTGGTGTAGCTAAGCGTGGAGCCATATGGAA 360
Db |
QY 361 GAAGCGGCTCCTGCTGACTCTGGCAAGATGATGATGAGTCTCTCAATACGCCATAGAGTC 420
Db |
QY 421 ATTGAAGAAGCCGTCTCAAAAGACTTCTGAACCTGGAGTGGTGTACTCCAGGCCAATATAAC 480
Db |
QY 481 ACACCTGCACAAATCGTATCTGCGAAGATGATGATGATGATGATGATGATGATGATGAT 540
Db |
QY 541 TTGCAAGAGCAGGTCGCAAGCGTTGATTTCTCTTAAGGTGTGATGATGATGATGATGAT 600
Db |
QY 601 GCTCTCTTGAAGCTGTAGCCAGAAATAGCTGAAATCTAGCTGAGTAAAGTTTTCAT 660
Db |
QY 661 GATTCTTACTGTCCTAGTGGCAATACAGAGCTGCTGTGATGATGATGATGATGATGAT 720
Db |
QY 721 GCTCAGCTCTTGAACGCTCAGGTCAAGGAACCCGTTGCTTCTATGAAAGTATTTGGGTC 780

Db 721 GCTCAGCTCTTGAACGCTCAGGTCAAGGAACCCGTTGCTTCTATGAAAGTATTTGGGTC 780
QY 781 ATGCAAGAGCAGGTCGCAAGCGTTGATTCGAGATGAGCCGGGAAAGTCTTGTTCAGGT 840
Db 781 ATGCAAGAGCAGGTCGCAAGCGTTGATTCGAGATGAGCCGGGAAAGTCTTGTTCAGGT 840
QY 841 TTGTTTAAATAAATGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGATTTA 900
Db 841 TTGTTTAAATAAATGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGATTTA 900
QY 901 GTAGCACTTTTAGAAAAATAG 921
Db 901 GTAGCACTTTTAGAAAAATAG 921
RESULT 2
AAH90730
ID AAH90730 standard; DNA; 921 BP.
XX
AC AAH90730;
XX 02-OCT-2001 (first entry)
XX CFE 34 coding sequence.
XX Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
XX CFE; CEG; Conserved Essential Gene; bacterial infection;
XX antisense therapy; antibiotic resistance; ds.
XX Streptococcus pneumoniae.
XX WC200149721-A2.
XX 12-JUL-2001.
XX 29-DEC-2000; 2000WO-US35604.
XX 30-DEC-1999; 99US-0174089.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE;
XX Thanassi JA;
XX WPI; 2001-496721/54.
XX P-PSDB; AAM01031.
PT Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX Claim 8; Pages 202-203; 380pp; English.
XX The present invention relates to nucleic acids (AAH90701-AAH90918)
XX encoding polypeptides (AAM01002-AAM01114), which are essential for the
XX viability of a bacterial cell wall. The acronym CFE stands for "Conserved
XX Expression", where CEG stands for "Conserved Essential Gene". The nucleic
XX acids are useful for detecting the presence of proteins essential for the
XX viability of a bacterial cell wall in samples such as cells, tissues,
XX biological fluids, blood, serum, nose, ear or throat swabs with ligands,
XX and for detecting corresponding target nucleic acid molecules with
XX complementary sequences. The nucleic acids are also useful for
XX determining whether a genomic nucleotide sequence of interest is
XX essential for viability of a bacterial cell or whether it resides within
XX an operon, by integrating an exogenous nucleotide sequence comprising a
XX portion of an open reading frame of the genomic sequence of interest
XX (comprising 200-500 base pairs) into the genomic sequence of interest
XX which confers a selectable phenotype to the cell, and determining cell
XX viability with a selection agent such as chloramphenicol. The nucleic
XX acids and proteins are also useful as vaccines and for treating bacterial
XX infections with gene therapy and antisense therapy. The nucleic acids
XX also enable identification of targets suitable for the treatment of

CC antibiotic resistant bacterial infections.

XX Sequence 921 BP; 246 A; 193 C; 232 G; 250 T; 0 other;

Query Match 99.1%; Score 913; DB 22; Length 921;

Best Local Similarity 99.5%; Pred. No. 1.1e-278;

Matches 916; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACTAAACAGCCCTTTTATTTGCTGTCAGAGTGCCCGAGTATCTAGGATGGGACGG 60
DB 1 ATGACTAAACAGCCCTTTTATTTGCTGTCAGAGTGCCCGAGTATCTAGGATGGGACGG 60
QY 61 GATTTCCTATGATCAGTATCCGATTCCTCAAGAAACGATTGATCGAGGAGTCAGGTGCTC 120
DB 61 GATTTCCTATGATCAGTATCCGATTCCTCAAGAAACGATTGATCGAGGAGTCAGGTGCTC 120
QY 121 GGTATGATTTACGTTATCTCATGATACGGAAGAACTCAATCAAGCCGCTAT 180
DB 121 GGTATGATTTACGTTATCTCATGATACGGAAGAACTCAATCAAGCCGCTAT 180
QY 181 ACGCAACAGCCATCTAGCGACTTCGGTTCCTATCTACCGTTTATTGCAAGAAAGGC 240
DB 181 ACGCAACAGCCATCTAGCGACTTCGGTTCCTATCTACCGTTTATTGCAAGAAAGGC 240
QY 241 TATCAGCCTGATATGTTGCTGGTTGCTCTTGAGAAATCTCTGCTTGGTGGCAAGC 300
DB 241 TATCAGCCTGATATGTTGCTGGTTGCTCTTGAGAAATCTCTGCTTGGTGGCAAGC 300
QY 301 GCGCCTTGGATTTGAGATGCGGTTCCTGCTAGCTAGCGTGGAGCCTATATGAA 360
DB 301 GCGCCTTGGATTTGAGATGCGGTTCCTGCTAGCTAGCGTGGAGCCTATATGAA 360
QY 361 GAAGCGGCTCCTGCTGACTCTGCGAAGATGTTAGCAGTTCTCAATAGCCAGTAGAGTC 420
DB 361 GAAGCGGCTCCTGCTGACTCTGCGAAGATGTTAGCAGTTCTCAATAGCCAGTAGAGTC 420
QY 421 ATTGAAGAGCCTGTCAAAAGCTTCTGAACTTGAGTGGTGTACTCCAGCCAACTATAAC 480
DB 421 ATTGAAGAGCCTGTCAAAAGCTTCTGAACTTGAGTGGTGTACTCCAGCCAACTATAAC 480
QY 481 ACACCTGCACAAATCGTCAATGCTGGAGAGTGGTTCAGTTGATCGAGCGGTGAACTT 540
DB 481 ACACCTGCACAAATCGTCAATGCTGGAGAGTGGTTCAGTTGATCGAGCGGTGAACTT 540
QY 541 TTGCAAGAGCAGGTGCGCAACCGTTGATTCCTCTTAAGGTGTCAGGTTCACACC 600
DB 541 TTGCAAGAGCAGGTGCGCAACCGTTGATTCCTCTTAAGGTGTCAGGTTCACACC 600
QY 601 GCTCTCCTTGAGCCTGTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAACTTTTCA 660
DB 601 GCTCTCCTTGAGCCTGTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAACTTTTCA 660
QY 661 GATTTTACTTGTCCCTAGTCGCAATACAGAACTGCTGTGATGCAAAAAGAGGACATT 720
DB 661 GATTTTACTTGTCCCTAGTCGCAATACAGAACTGCTGTGATGCAAAAAGAGGACATT 720
QY 721 GCTAGCTCTTGAGCGTCAGGTCAAGAACCCGTTGTTCTATGAAAGTATGGGTC 780
DB 721 GCTAGCTCTTGAGCGTCAGGTCAAGAACCCGTTGTTCTATGAAAGTATGGGTC 780
QY 781 ATGCAAGAGCAGGCATAGCAACTTTATCAGATTTGGACCGGGAAAGTCTGTCAAGT 840
DB 781 ATGCAAGAGCAGGCATAGCAACTTTATCAGATTTGGACCGGGAAAGTCTGTCAAGT 840
QY 841 TTTGTTAAAAAATTGATCAAACTGCTCATTAGCTCATGTGGAAGTCAAGCGAGTTTA 900
DB 841 TTTGTTAAAAAATTGATCAAACTGCTCATTAGCTCATGTGGAAGTCAAGCGAGTTTA 900
QY 901 GTAGCATTTTAGAAAAATAG 921
DB 901 GTAGCATTTTAGAAAAATAG 921

RESULT 3

AAV52140/C
AAV52140 standard; DNA; 19702 BP.

XX AAV52140;
AC AAV52140;

DT 23-OCT-1998 (first entry)

XX Streptococcus pneumoniae genome fragment SEQ ID NO:7.

XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
computer readable medium; vaccine; pharmaceutical composition; ds.

OS Streptococcus pneumoniae.

PN WO9818931-A2.

XX 07-MAY-1998.

PF 30-OCT-1997; 97WO-US19588.

PR 31-OCT-1996; 96US-0029960.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Barash SC, Choi GH, Dillion PJ, Dougherty BA, Fannon M;
PI Kunsch CA, Rosen CA;

DR WPI; 1998-272225/24.

XX Computer-readable medium with recorded Streptococcus pneumoniae
polynucleotide sequences - useful in diagnostic kits and assays, and
pharmaceutical compositions and vaccines for Streptococcus
pneumoniae

PS Claim 1; Page 194-205; 1409pp; English.

XX The present invention describes a computer readable medium which has
the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
recorded on it, or a representative fragment or a sequence at least 95%
identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
Streptococcus pneumoniae. The present invention also describes an
isolated nucleic acid molecule encoding a homologue of any of the
fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
nucleic acid molecule is produced by a process comprising: (a) screening
a genomic DNA library using as a probe a target sequence defined by any
of the sequences in SEQ ID NO:1 to 391, identifying members of the
library which contain sequences that hybridize to the target sequence and
isolating the nucleic acid molecules from the members; or (b) isolating
mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
molecules whose nucleotide sequence is homologous to amplification
primers derived from the fragment of the S. pneumoniae genome to prime
the amplification and isolating the amplified sequences. The computer
readable medium can be used in a computer-based system for identifying
fragments of the S. pneumoniae genome of commercial importance, or
expression modulating fragments of the S. pneumoniae genome. Products
from the present invention can be used in diagnosis kits and assays, and
pharmaceutical compositions and vaccines for S. pneumoniae.

XX Sequence 19702 BP; 5736 A; 4548 C; 3728 G; 5690 T; 0 other;

Query Match 99.1%; Score 913; DB 19; Length 19702;

Best Local Similarity 99.5%; Pred. No. 5.8e-278;

Matches 916; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGACTAAACAGCCCTTTTATTTGCTGTCAGAGTGCCCGAGTATCTAGGATGGGACGG 60

DB 17149 ATGACTAAACAGCCCTTTTATTTGCTGTCAGAGTGCCCGAGTATCTAGGATGGGACGG 17090

QY 61 GATTTCTATGATCAGTATCCGATTTGTCAGAAACGATGTCAGCGAGTCAGGTGCTC 120

DB 17089 GATTTCTATGATCAGTATCCGATTTGTCAGAAACGATGTCAGCGAGTCAGGTGCTC 17030

```
QY 121 GGTATGATTTACGTTATCTCATCGATACGGAAGAGACAAACTCAATCAGACCCGCTAT 180
Db |||||||
QY 17029 GGTATGATTTACGTTATCTCATCGATACGGAAGAGACAAACTCAATCAGACCCGCTAT 16970
Db |||||||
QY 181 ACGCAACCGACCATTCCTAGGACATTCGGTTCGTTATCTACCGTTTATTCGAAGAAAAGGCG 240
Db |||||||
QY 16969 ACGCAACCGACCATTCCTAGGACATTCGGTTCGTTATCTACCGTTTATTCGAAGAAAAGGCG 16910
Db |||||||
QY 241 TATCAGCCTGATATGTTGCTGTTGCTCTCTGGAGATACCTCTGCTTGGTGGCAAGC 300
Db |||||||
QY 16909 TATCAGCCTGATATGTTGCTGTTGCTCTCTGGAGATACCTCTGCTTGGTGGCAAGT 16850
Db |||||||
QY 301 GGGCCCTTGAATTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGGAGCCTATATGGA 360
Db |||||||
QY 16849 GGGCCCTTGAATTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGGAGCCTATATGGA 16790
Db |||||||
QY 361 GAAGCGCTCTGCTGATCTCTGCAAGATGTTAGAGTTCCTCAATACGCGAGTAGAGTTC 420
Db |||||||
QY 16789 GAAGCGCTCTGCTGATCTCTGCAAGATGTTAGAGTTCCTCAATACGCGAGTAGAGTTC 16730
Db |||||||
QY 421 ATTGAAGAAGCCTGTCAAAAAGCTTCTGAACTTGGAGTGGTTACTCCAGCCAACTATAAC 480
Db |||||||
QY 16729 ATTGAAGAAGCCTGTCAAAAAGCTTCTGAACTTGGAGTGGTTACTCCAGCCAACTATAAC 16670
Db |||||||
QY 481 ACACCTGCAAAATCGTCATTTGCTGGAGAGTGGTTGCAAGTTCAGTTCAGCGGTTGAAC 540
Db |||||||
QY 16669 ACACCTGCAAAATCGTCATTTGCTGGAGAGTGGTTGCAAGTTCAGTTCAGCGGTTGAAC 16610
Db |||||||
QY 541 TTGCAAGAAGCAGGTGCAAAAGCCTGATTCCTTAAAGTGTGAGTCCCTCCCTTCACACC 600
Db |||||||
QY 16609 TTGCAAGAAGCAGGTGCAAAAGCCTGATTCCTTAAAGTGTGAGTCCCTCCCTTCACACC 16550
Db |||||||
QY 601 GCTCTCTTGAGCCTGTAGCCGAAACTAGTCTGAACTCTAGCTCAGTAAAGTTTTC 660
Db |||||||
QY 16549 GCTCTCTTGAGCCTGTAGCCGAAACTAGTCTGAACTCTAGCTCAGTAAAGTTTTC 16490
Db |||||||
QY 661 GATTTACTTGTCCTTGTGCGCAATACAGAGTCTGCTGATGCAAAAGAGGACATT 720
Db |||||||
QY 16489 GATTTACTTGTCCTTGTGCGCAATACAGAGTCTGCTGATGCAAAAGAGGACATT 16430
Db |||||||
QY 721 GCTCAGCTCTGAGCGTCAAGTCAAGGAAACCGTTTCGTTTCTATGAAAGTATGGGTC 780
Db |||||||
QY 16429 GCTCAGCTCTGAGCGTCAAGTCAAGGAAACCGTTTCGTTTCTATGAAAGTATGGGTC 16370
Db |||||||
QY 781 ATGCAAGAAGCAGCATACGCACTTTATCGAGATTGGACCGGGAAGCTTTGTCAGT 840
Db |||||||
QY 16369 ATGCAAGAAGCAGCATACGCACTTTATCGAGATTGGACCGGGAAGCTTTGTCAGT 16310
Db |||||||
QY 841 TTGTTAAAAAATTGATCAAACTGCTCACTTACGCTCATGTGGAAGATCAAGCGAGTTTA 900
Db |||||||
QY 16309 TTGTTAAAAAATTGATCAAACTGCTCACTTACGCTCATGTGGAAGATCAAGCGAGTTTA 16250
Db |||||||
QY 901 GTAGCACTTTAGAAAAATAG 921
Db |||||||
QY 16249 GTAGCACTTTAGAAAAATAG 16229
Db |||||||
RESULT 4
ID AAS55845 standard; DNA; 921 BP.
XX AAS55845;
AC AAS55845;
XX
DT 13-FEB-2002 (first entry)
XX
DE Streptococcus pneumoniae DNA for cellular proliferation protein #416.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene;
XX antibiotic; antibacterial; drug design.
XX
OS Streptococcus pneumoniae.
XX
```

WO200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207727P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-269308P.

(ELIT-) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI; 2001-611495/70.

P-PsDB; AAU37986.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -

Claim 27; Seq ID No 9482; 511pp; English.

The invention relates to antisense inhibitors of genes essential to

prokaryotic cellular proliferation, their use in identifying the

genes their use in the discovery of novel antibiotics, the essential

genes themselves and the encoded proteins. The prokaryotes used are

Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

invention is also useful for the identification of potential new targets

for antibiotic development. The antisense nucleic acids can also be used

to identify proteins used in proliferation, to express these proteins,

and to obtain antibodies capable of binding to the expressed proteins.

The proteins can be used to screen compounds in rational drug discovery

programmes. The antisense nucleic acid sequence is also useful to screen

for homologous nucleic acids which are required for cell proliferation in

a wide variety of organisms. The present sequence encodes an

essential prokaryotic cellular proliferation protein.

Note: The sequence data for this patent did not form part

of the printed specification, but was obtained in electronic

format directly from WIPO at

ftp.wipo.int/pub/published_pct_sequences.

Sequence 921 BP; 246 A; 191 C; 232 G; 252 T; 0 other;

Query Match 99.0%; Score 911.4; DB 23; Length 921;

Best Local Similarity 99.3%; Pred No 3.5e-278;

Matches 915; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 661 GATTTTACTTGTCCCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 720
D5 1378 GATTTTACTTGTCCCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 1437
QY 721 GCTCAGCTCTTGGCGCTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 780
D5 1438 GCTCAGCTCTTGGCGCTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 1497
QY 781 ATGCAAGAAGCAGGTCATAGCACTTTATCGAGATTGGACCGGGGAAAGTCTTGTTCAGGT 840
D5 1498 ATGCAAGAAGCAGGTCATAGCACTTTATCGAGATTGGACCGGGGAAAGTCTTGTTCAGGT 1557
QY 841 TTTGTTAAAAAATTGATCAAACTGCTCACTTATAGCTCATGTGGAAGATCAAGCGAGTTTA 900
D5 1558 TTTGTTAAAAAATTGATCAAACTGCTCACTTATAGCTCATGTGGAAGATCAAGCGAGTTTA 1617
QY 901 GTAGCACTTTTAAAAAATAG 921
D5 1618 GTAGCACTTTTAAAAAATAG 1638

RESULT 6

AAH90843
ID AAH90843 standard; DNA; 945 BP.
XX
AC AAH90843;
XX
DT 02-OCT-2001 (first entry)
XX
DE 2CPE 34 coding sequence.
XX
KW Antibacterial; vaccine; gene therapy; bacterial cell wall viability;
KW CPE; CEG; Conserved Essential Gene; bacterial infection;
KW antisense therapy; antibiotic resistance; ds.
XX
OS Streptococcus pneumoniae.
XX
PN WO200149721-A2.
XX
PD 12-JUL-2001.
XX
PF 29-DEC-2000; 2000WO-US35604.
XX
PR 30-DEC-1999; 99US-0174089.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Dougherty TJ, Pucci MJ, Dougherty BA, Davison DB, Bruccoleri RE,
PI Thanassi JA;
XX
DR WPI; 2001-496721/54.
XX
XX
XX Nucleic acids encoding conserved essential genes involved in bacterial
PT replication which are potential targets for the treatment of antibiotic
PT resistant bacterial infections -
XX
PS Claim 8; Fig 55; 380pp; English.
XX
XX The present invention relates to nucleic acids (AAH90701-AAH90918)
CC encoding polypeptides (AAH01002-AAH01114), which are essential for the
CC viability of a bacterial cell wall. The acronym CPE stands for "CEG For
CC Expression", where CEG stands for "Conserved Essential Gene". The nucleic
CC acids are useful for detecting the presence of proteins essential for the
CC viability of a bacterial cell wall in samples such as cells, tissues,
CC biological fluids, blood, serum, nose, ear or throat swabs with ligands,
CC and for detecting corresponding target nucleic acid molecules with
CC complementary sequences. The nucleic acids are also useful for
CC determining whether a genomic nucleotide sequence of interest is
CC essential for viability of a bacterial cell or whether it resides within
CC an operon, by integrating an exogenous nucleotide sequence comprising a
CC portion of an open reading frame of the genomic sequence of interest
CC (comprising 200-500 base pairs) into the genomic sequence of interest
CC which confers a selectable phenotype to the cell, and determining cell

CC viability with a selection agent such as chloramphenicol. The nucleic
CC acids and proteins are also useful as vaccines and for treating bacterial
CC infections with gene therapy and antisense therapy. The nucleic acids
CC also enable identification of targets suitable for the treatment of
CC antibiotic resistant bacterial infections.
XX
SQ Sequence 945 BP; 254 A; 205 C; 232 G; 254 T; 0 other;
Query Match 98.6%; Score 908.4; DB 22; Length 945;
Best Local Similarity 99.3%; Pred. No. 3.1e-277; Indels 0; Gaps 0;
Matches 912; Conservative 0; Mismatches 6;
QY 1 ATGACTAAACAGCCTTTTATTTGCTGTCAGAGTCCCGAGTATCTAGGATGGGACGG 60
D5 1 ATGACTAAACAGCCTTTTATTTGCTGTCAGAGTCCCGAGTATCTAGGATGGGACGG 60
QY 61 GATTTCTATGATCAGTATCGATTGTCAAGAAACGATTGATCGAGGAGTCAGGTCTC 120
D5 61 GATTTCTATGATCAGTATCGATTGTCAAGAAACGATTGATCGAGGAGTCAGGTCTC 120
QY 121 GGTATGATTTACGTTATCTCATCGATACGAAAGACAAACTCAATCAGACCCGCTAT 180
D5 121 GGTATGATTTGCGTTATCTCATCGATACGAAAGACAAACTCAATCAGACCCGCTAT 180
QY 181 ACGCAACAGCCATTCTAGGACTTGGTGGTCTTATACCGTTTATTTGCAAGAAAGGCG 240
D5 181 ACGCAACAGCCATTCTAGGACTTGGTGGTCTTATACCGTTTATTTGCAAGAAAGGCG 240
QY 241 TATCAGCCTGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
D5 241 TATCAGCCTGATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 300
QY 301 GCGCCTTGGATTGTTGAAGATGCGTTGCTTGGTAGCTAAGCGTGGAGCTATATGGAA 360
D5 301 GCGCCTTGGATTGTTGAAGATGCGTTGCTTGGTAGCTAAGCGTGGAGCTATATGGAA 360
QY 361 GAAGCGGCTCTGCTGACTCTGCAAGATGTTAGCAGTCTCTCAATACGCCAGTAGAGTTC 420
D5 361 GAAGCGGCTCTGCTGACTCTGCAAGATGTTAGCAGTCTCTCAATACGCCAGTAGAGTTC 420
QY 421 ATTGAAGAAGCTGTCAAAAAGCTTCTGAACCTTGGAGTGGTGTACTCCAGCAACTATAAC 480
D5 421 ATTGAAGAAGCTGTCAAAAAGCTTCTGAACCTTGGAGTGGTGTACTCCAGCAACTATAAC 480
QY 481 ACACCTGCACAAATCGTCATTGCTGGAGAAAGTGGTTCAGTTCATCGAGCGTTGAACCT 540
D5 481 ACACCTGCACAAATCGTCATTGCTGGAGAAAGTGGTTCAGTTCATCGAGCGTTGAACCT 540
QY 541 TTGCAAGAAGCAGGTGCCAAACGCTTGATTCCTTTAAGGTGTCAGTCCCTTTCAACC 600
D5 541 TTGCAAGAAGCAGGTGCCAAACGCTTGATTCCTTTAAGGTGTCAGTCCCTTTCAACC 600
QY 601 GCTCTCCTTGGAGCTGCTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAACTTTTCA 660
D5 601 GCTCTCCTTGGAGCTGCTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAACTTTTCA 660
QY 601 TCTCTCTTGAACCTGCTAGCCAGAACTAGCTGAACTCTGCTCAGGTAACTTTTCA 660
D5 601 TCTCTCTTGAACCTGCTAGCCAGAACTAGCTGAACTCTGCTCAGGTAACTTTTCA 660
QY 661 GATTTTACTTGTCCCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 720
D5 661 GATTTTACTTGTCCCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 720
QY 721 GCTCAGCTCTTGGCGCTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 780
D5 721 GCTCAGCTCTTGGCGCTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 780
QY 781 ATGCAAGAAGCAGGTCATAGCACTTTATCGAGATTGGACCGGGGAAAGTCTTGTTCAGGT 840
D5 781 ATGCAAGAAGCAGGTCATAGCACTTTATCGAGATTGGACCGGGGAAAGTCTTGTTCAGGT 840
QY 841 TTTGTTAAAAAATTGATCAAACTGCTCACTTATAGCTCATGTGGAAGATCAAGCGAGTTTA 900
D5 841 TTTGTTAAAAAATTGATCAAACTGCTCACTTATAGCTCATGTGGAAGATCAAGCGAGTTTA 900
QY 901 GTAGCACTTTTAAAAAATAG 918

Db 901 GTAGCATTGTAGAAAA 918
 RESULT 7
 AAV37241
 ID AAV37241 standard; DNA; 547 BP.
 XX
 AC AAV37241;
 XX
 DT 24-SEP-1998 (first entry)
 XX
 DE C-terminal coding region encoding a FabD polypeptide embodiment.
 XX
 KW FabD protein; malonyl-CoA:ACP family; diagnosis; infection; vaccine;
 KW screen; ss.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN WO9822133-A1.
 XX
 PD 28-MAY-1998.
 XX
 PF 14-NOV-1997; 97WO-US20992.
 XX
 PR 18-NOV-1996; 95US-0031160.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX
 PI Gentry DR, Lonsdale JT, Payne DJ, Pearson SC, Van Aller G;
 XX
 DR WPI; 1998-312173/27.
 DR P-PSDB; AAW60858.
 XX
 PT New isolated Streptococcus pneumoniae FabD gene - used to develop
 PT products for the diagnosis, prevention and treatment of bacterial
 PT diseases, particularly S. pneumoniae infection
 PS
 PS Disclosure; Pages 6-7; 45pp; English.
 XX
 CC The present sequence represents the C-terminal coding region of FabD
 CC embodiment (coding begins at nucleotide 2) of Streptococcus pneumoniae
 CC 010093 (NCIMB 40800). The novel FabD polypeptides are related to other
 CC proteins of the malonyl-CoA:ACP family. The products can be used for the
 CC diagnosis of Streptococcus pneumoniae infections. Vectors containing
 CC the FabD DNA sequence can be administered directly to a mammal to
 CC produce the FabD peptide to provoke an antibody/i-cell response in
 CC order to prevent a disease. The peptide can be used to screen for
 CC compounds which modulate its activity.
 XX
 SQ Sequence 547 BP; 154 A; 115 C; 132 G; 146 T; 0 other;
 Query Match 57.5%; Score 529.4; DB 19; Length 547;
 Best Local Similarity 98.0%; Pred. No. 3.2e-157;
 Matches 536; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 372 TGCTGACTCTGCGAAGATGTTAGCAGTTCTCAATACGCCAGTAGAGTTCATTGAAGAAGC 431
 Db 1 TGTGTAAGTTGCGAAGATGTTAGCAGTTCTCAATACGCCAGTAGAGTTCATTGAAGAAGC 60
 QY 432 CTGTCAAAAAGCTTCTGAACTTGGAGTGGTGTACTCCAGCCGCACTATACACACTGTCACA 491
 Db 61 CTGTCAAAAAGCTTCTGAACTTGGAGTGGTGTACTCCAGCCGCACTATACACACTGTCACA 120
 QY 492 AATCGTCAATTCCTGGAGAAGTGGTTGCAGTTGATCGAGCGGTTGAAGTTTGAAGAAGC 551
 Db 121 AATCGTCAATTCCTGGAGAAGTGGTTGCAGTTGATCGAGCGGTTGAAGTTTGAAGAAGC 180
 QY 552 AGGTGCCAAAGCTTGTATCTCTTAAGGTGTACAGTCCCTTTACACCGCTCTCCTTGA 611
 Db 181 AGGTGCCAAAGCTTGTATCTCTTAAGGTGTACAGTCCCTTTACACCGCTCTCCTTGA 240
 QY 612 CCCTGCTAGCCAGAAAGCTAGCTGAACTCTAGCTCAGGTAAGTTTTCAGATTTTACTTG 671

Db 241 GCCTGCTAGCCAGAACTAGCTGAACCTAGCTCAGTAACTTTTTCAGATTTTACTTG 300
 QY 672 TCCCTAGTCGGCAATACAGAAAGCTGCTGTATGCAAAAAGAGACATTGCTCAGCTCTT 731
 Db 301 TCCCTAGTCGGCAATACAGAAAGCTGCTGTATGCAAAAAGAGACATTGCTCAGCTCTT 360
 QY 732 GACGCGTCAGTCAAGGAACCCGTTTCGTTTCTATGAAAGTATTGGGGTTCATGCAAGAAGC 791
 Db 361 GACGCGTCAGTCAAGGAACCCGTTTCGTTTCTATGAAAGTATTGGGGTTCATGCAAGAAGC 420
 QY 792 AGCATATAGCAAACTTTATCGAGATTGGACCGGGGAAAGTCTTGTGAGGTTTGTAAAAA 851
 Db 421 AGCATATAGCAAACTTTATCGAGATTGGACCGGGGAAAGTCTTGTGAGGTTTGTAAAAA 480
 QY 852 AATTGATCAAACTGCTCAGTTCAGTTCAGTCAATGGAAGATCAAGCGAGTTAGTAGCAGCTTTT 911
 Db 481 AATTGATCAAACTGCTCAGTTCAGTTCAGTCAATGGAAGATCAAGCGAGTTAGTAGCAGCTTTT 540
 QY 912 AGAAAAA 918
 Db 541 AGAAAAA 547
 RESULT 8
 AAT98542
 ID AAT98542 standard; DNA; 1196 BP.
 XX
 AC AAT98542;
 XX
 DT 06-NOV-1998 (first entry)
 XX
 DE DNA encoding a S. pneumoniae 3-oxoacyl reductase precursor protein.
 XX
 KW Streptococcus pneumoniae protein; genetic immunisation; antagonist;
 KW immunological response; inoculation; antibody production; inhibitor;
 KW T cell immune response; antimicrobial compound; bacterial adhesion;
 KW extracellular matrix protein; protein-mediated cell invasion; wound;
 KW pathogenesis; ss.
 XX
 OS Streptococcus pneumoniae.
 XX
 FH Key Location/Qualifiers
 FT CDS 586..1146
 FT /*tag= a
 XX
 PN WO9743303-A1.
 XX
 PD 20-NOV-1997.
 XX
 PF 14-MAY-1997; 97WO-US07950.
 XX
 PR 14-MAY-1996; 95US-0017670.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
 PI Stodola RK;
 XX
 DR WPI; 1998-008793/01.
 DR P-PSDB; AAW38474.
 XX
 PT Novel Streptococcus pneumoniae proteins and related DNA - useful for
 PT diagnosing anti-microbial agents for treatment of bacterial
 PT infections
 XX
 PS Claim 4; Page 90; 483pp; English.
 XX
 CC This sequence encodes a Streptococcus pneumoniae protein that (based on
 CC homology with a Cuphea lanceolata protein) is a 3-oxoacyl reductase
 CC precursor protein, and represents a DNA sequence of the invention.
 CC The DNA sequences were isolated from Streptococcus pneumoniae strain

CC 0100993 (NCIMB 40794). The Streptococcus pneumoniae proteins of the
 CC invention can be used to identify compounds which interact with and
 CC inhibit or activate the activity of the proteins. Antagonists can be
 CC used to treat diseases caused by S. pneumoniae proteins, through genetic
 CC immunisation. They can also be used to induce an immunological response
 CC in a mammal by inoculation with the S. pneumoniae proteins or delivery
 CC of the encoding nucleic acids in a vector adequate to produce antibody
 CC and/or T cell immune responses to protect the animal from disease. The
 CC proteins can also be used to identify antimicrobial compounds which are
 CC capable of inhibiting their bioactivity. In particular the proteins of
 CC the invention can be used to prevent adhesion of bacteria to mammalian
 CC extracellular matrix proteins on in-dwelling devices or in wounds, to
 CC block protein-mediated mammalian cell invasion, and to block the normal
 CC progression of pathogenesis in infections initiated other than by the
 CC implantation of in-dwelling devices or other surgical techniques.
 XX
 XX Sequence 1199 BP; 355 A; 223 C; 286 G; 332 T; 0 other;

Query Match 56.8%; Score 523.2; DB 19; Length 1196;
 Best Local Similarity 99.1%; Pred. No. 4.6e-155;
 Matches 547; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
 QY 372 TGTGACTCTGCAAGATGCTAGCAGTTCCTCAATACGCCAGTACGAGTCAATTGAGAAGC 431
 Db 1 TGTGAACTTGGCAAGATGCTAGCAGTTCCTCAATACGCCAGTACGAGTCAATTGAGAAGC 60
 QY 432 CTGTCAAAAGCTTCT-GAAGTGGAGTGGTGTACTCCAGCCAACTATACACACCTGCAC 490
 Db 61 CTGTCAAAAGCTTCTGGAACTTGGAGTGGTGTACTCCAGCCAACTATACACACCTGCAC 120
 QY 491 AAATCGTCAATTCGTGAGAAAGTGGTTCGATTCGAGCGGTGAACTTTTGAAGAAG 550
 Db 121 AAATCGTCAATTCGTGAGAAAGTGGTTCGATTCGAGCGGTGAACTTTTGAAGAAG 180
 QY 551 CAGTCCCAAGCTTGAATTCCTCTTAAGTGTCTAGGTCGAGTCCCTTTCACACCGCTCTCCTTG 610
 Db 181 CAGTCCCAAGCTTGAATTCCTCTTAAGTGTCTAGGTCGAGTCCCTTTCACACCGCTCTCCTTG 240
 QY 611 AGCTCTGCTAGCCAGAACTAGTGAAGTCTAGCTCAGTAAAGTCTTTTTCAGATTTTACTT 670
 Db 241 AGCTCTGCTAGCCAGAACTAGTGAAGTCTAGCTCAGTAAAGTCTTTTTCAGATTTTACTT 300
 QY 671 GTCCCTCTAGTGGCAATACAGAGCTGTGTATGCAAAAGAGACATTTGCTAGCTCT 730
 Db 301 GTCCCTCTAGTGGCAATACAGAGCTGTGTATGCAAAAGAGACATTTGCTAGCTCT 360
 QY 731 TGACGGTTCAGGTCAGGAACCGGTTTCGTTTCTATGAAAGTATTCGGGTTCATGCAAGAAG 790
 Db 361 TGACGGTTCAGGTCAGGAACCGGTTTCGTTTCTATGAAAGTATTCGGGTTCATGCAAGAAG 420
 QY 791 CAGGCATAAGCAACTTTA-TCGAGATTTGACCGGGGAAAGCTTTGTCAGGTTTGTATAA 849
 Db 421 CAGGCATAAGCAACTTTA-TCGAGATTTGACCGGGGAAAGCTTTGTCAGGTTTGTATAA 480
 QY 850 AAAATTGATCAACTCTCCTCACTTAGCTCATGTGGAAGTCAAGCAGTTTACTAGCACTT 909
 Db 481 AAAATTGATCAACTCTCCTCACTTAGCTCATGTGGAAGTCAAGCAGTTTACTAGCACTT 540
 QY 910 TTAGAAAAATAG 921
 Db 541 TTAGAAAAATAG 552

RESULT 9
 AAX30801/c
 ID AAX30801 standard; DNA; 1199 BP.
 XX
 AC AAX30801;
 XX
 DT 20-MAY-1999 (first entry)
 XX
 DE Streptococcus pneumoniae genomic DNA sequence SEQ ID NO:78.
 XX

KW Streptococcus pneumoniae strain 0100993; vaccine; immune response;
 KW streptococcal infection; pneumococcal; ss.
 XX
 OS Streptococcus pneumoniae.
 PN W09737026-A1.
 PD 09-OCT-1997.
 XX
 XX 01-APR-1997; 97WO-US05306.
 PR 22-AUG-1996; 96US-0025788.
 PR 02-APR-1996; 96US-0014690.
 XX
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 XX Black MT, Hodgson JE, Knowles DJC, Nicholas RO;
 PI Stodola RK;
 XX
 XX WPI; 1997-503111/46.
 DR P-PSDB; AAY11212.
 XX
 XX Nucleic acids encoding pneumococcal polypeptide(s) - useful in
 PT vaccines, drug screening, etc
 XX
 XX Claim 5; Page 105-106; 354pp; English.
 PS
 XX AAX30724 to AAX30946 represent genomic DNA sequences isolated from
 CC Streptococcus pneumoniae strain 0100993. These genomic DNA sequences
 CC encode the novel proteins given in AAY1114 to AAY11367. The proteins,
 CC isolated from Streptococcus pneumoniae, can be used in vaccines against
 CC streptococcal infections and in assays for identifying compounds that
 CC inhibit or activate the activity of the proteins. The antagonists can
 CC be used to treat an individual having need to inhibit a bacterial
 CC protein. Vectors expressing the proteins can be used to induce a
 CC protective immune response in mammals.
 XX
 XX Sequence 1199 BP; 334 A; 287 C; 223 G; 355 T; 0 other;

Query Match 56.8%; Score 523.2; DB 18; Length 1199;
 Best Local Similarity 99.1%; Pred. No. 4.6e-155;
 Matches 547; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
 QY 372 TGTGACTCTGCAAGATGCTAGCAGTTCCTCAATACGCCAGTACGAGTCAATTGAGAAGC 431
 Db 1199 TGTGAACTTGGCAAGATGCTAGCAGTTCCTCAATACGCCAGTACGAGTCAATTGAGAAGC 1140
 QY 432 CTGTCAAAAGCTTCT-GAAGTGGAGTGGTGTACTCCAGCCAACTATACACACCTGCAC 490
 Db 1139 CTGTCAAAAGCTTCTGGAACTTGGAGTGGTGTACTCCAGCCAACTATACACACCTGCAC 1080
 QY 491 AAATCGTCAATTCGTGAGAAAGTGGTTCGAGTTCGATCGAGCGGTGAACTTTTGAAGAAG 550
 Db 1079 AAATCGTCAATTCGTGAGAAAGTGGTTCGAGTTCGATCGAGCGGTGAACTTTTGAAGAAG 1020
 QY 551 CAGTCCCAAGCTTGAATTCCTCTTAAGTGTCTAGGTCGAGTCCCTTTCACACCGCTCTCCTTG 610
 Db 1019 CAGTCCCAAGCTTGAATTCCTCTTAAGTGTCTAGGTCGAGTCCCTTTCACACCGCTCTCCTTG 960
 QY 611 AGCTCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGTAAAGTCTTTTTCAGATTTTACTT 670
 Db 959 AGCTCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGTAAAGTCTTTTTCAGATTTTACTT 900
 QY 671 GTCCCTCTAGTGGCAATACAGAGCTGTGTATGCAAAAGAGGACATTTGCTAGCTCT 730
 Db 899 GTCCCTCTAGTGGCAATACAGAGCTGTGTATGCAAAAGAGGACATTTGCTAGCTCT 840
 QY 731 TGACGGTTCAGGTCAGGAACCGGTTTCGTTTCTATGAAAGTATTCGGGTTCATGCAAGAAG 790
 Db 839 TGACGGTTCAGGTCAGGAACCGGTTTCGTTTCTATGAAAGTATTCGGGTTCATGCAAGAAG 780
 QY 791 CAGGCATAAGCAACTTTA-TCGAGATTTGACCGGGGAAAGCTTTGTCAGGTTTGTATAA 849

779 CAGCATAGCACTTTATTCGAGATTGGACCGGGGAAAGCTTGTGAGGTTTGTAAA 720
 850 AAAATTGATCAAACTGCTACTAGCTCATGTGGAAGATCAACGAGTTTGTAGCACTT 909
 719 AAAATTGATCAAACTGCTACTAGCTCATGTGGAAGATCAACGAGTTTGTAGCACTT 660
 910 TTAGAAAAAATAG 921
 659 TTAGAAAAAATAG 648

RESULT 10
 AAV37406/C
 ID AAV37406 standard; DNA; 1209 BP.
 AC AAV37406;
 XX
 DT 13-OCT-1998 (first entry)
 DE Streptococcus pneumoniae coding region.
 XX
 KW coding region; ORF; open reading frame; antibacterial;
 KW infection; prevention; meningitis; ss.
 XX
 OS Streptococcus pneumoniae.
 XX
 FH Key Location/Qualifiers
 CDS complement (731..1123)
 FT /*tag= a
 FT /product= putative malonyl coenzyme A-acyl carrier
 FT protein transacylase homologue

WO9819689-A1.
 XX
 XX 14-MAY-1998.
 XX
 PF 27-OCT-1997; 97WO-US19226.
 XX
 PR 01-NOV-1996; 96US-0029930.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Black MT, Hodgson JE, Knowles DJC, Lonetto MA, Nicholas RO;
 PI Reid RH, Zarfos PN;
 XX
 DR WPI; 1998-286586/25.
 DR P-PSDB; AAW61018.
 XX

XX
 PT New isolated nucleic acids from Streptococcus pneumoniae - useful,
 FT e.g. for identifying anti-bacterial(s) for treatment and prevention
 FT of meningitis
 XX
 PS Claim 1; Page 111-112; 130pp; English.

XX
 CC The sequence is that of a coding region isolated from
 CC S. pneumoniae. Its encoded protein, or agonists of it,
 CC may be useful as an antibacterial for treatment or
 CC prevention of infection, specifically caused by S. pneumoniae
 CC (particularly meningitis) but possibly also Helicobacter
 CC pylori (ulcers and gastric cancer). It may be of particular
 CC use before insertion of an in-dwelling device or any other
 CC invasive procedure. The protein, or nucleic acid encoding
 CC it, can also be used in vaccines to induce a cellular
 CC and/or humoral immune response, or to screen for other
 CC antibacterials. The DNA may also contain flanking sequences
 CC that are potential sources of control elements for bacterial
 CC gene expression. Detecting a sequence encoding the protein
 CC can be used diagnostically, e.g. to detect a mutation for
 CC serotyping or classifying infectious agents.

XX
 SQ Sequence 1209 BP; 336 A; 288 C; 227 G; 358 T; 0 other;

Query Match 56.8%; Score 523.2; DB 19; Length 1209;
 Best Local Similarity 99.1%; Pred. No. 4.6e-155;
 Matches 547; Conservative 0; Mismatches 3; Indels 2; Gaps 2;
 QY 372 TGCTGACTCTGGCAAGATGGTAGGAGTTCTCAATACGCCAGTAGAGGTCAATTGAAGAAGC 431
 DB 1209 TGCTGAACTTGGCAAGATGGTAGGAGTTCTCAATACGCCAGTAGAGGTCAATTGAAGAAGC 1150
 QY 432 CTGTCAAAAAGCTTCT-GAACTTGGAGTGGTTACTTCCAGCCCAACTATTAACACACTGCAC 490
 DB 1149 CTGTCAAAAAGCTTCTGGAACTTGGAGTGGTTACTTCCAGCCCAACTATTAACACACTGCAC 1090
 QY 491 AAATCGTCAATTGCTGGAGAGTGGTTGCAGTTGATCGAGCGTTGAAGCTTTTGAAGAAG 550
 DB 1089 AAATCGTCAATTGCTGGAGAGTGGTTGCAGTTGATCGAGCGTTGAAGCTTTTGAAGAAG 1030
 QY 551 CAGGTGCCAAAAGCTTGTATTCCTCTTAAAGGTGCAGGTCCCTTTTACACACCCTCTCCCTTG 610
 DB 1029 CAGGTGCCAAAAGCTTGTATTCCTCTTAAAGGTGCAGGTCCCTTTTACACACCCTCTCCCTTG 970
 QY 611 AGCTGTAGTCAGCAAACTAGCTGAAACTCTAGCTCAGGTAAGTTTTTTCAGATTTTACTT 670
 DB 969 AGCTGTAGTCAGCAAACTAGCTGAAACTCTAGCTCAGGTAAGTTTTTTCAGATTTTACTT 910
 QY 671 GTCCCTAGTCGGCAATACAGAACTGCTGTGATGCAAAAAGAGAGACATTGCTCAGCTCT 730
 DB 909 GTCCCTAGTCGGCAATACAGAACTGCTGTGATGCAAAAAGAGAGACATTGCTCAGCTCT 850
 QY 731 TGACGCTCAGTCAAGGAACCCGTTCTGTTCTATGAAGTATTGGGGTCAATGAAGAAG 790
 DB 849 TGACGCTCAGTCAAGGAACCCGTTCTGTTCTATGAAGTATTGGGGTCAATGAAGAAG 790
 QY 791 CAGGCATAAGCAACTTTA-TGAGATTGGACCGGGGAAAGTCTTGTGAGGTTTGTGTTAAA 849
 DB 789 CAGGCATAAGCAACTTTATTCGAGATTGGACCGGGGAAAGTCTTGTGAGGTTTGTGTTAAA 730
 QY 850 AAAATTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGCACTT 909
 DB 729 AAAATTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGCACTT 670
 QY 910 TTAGAAAAAATAG 921
 DB 669 TTAGAAAAAATAG 658

RESULT 11
 ID ABN68641
 ID ABN68641 standard; DNA; 942 BP.
 XX
 AC ABN68641;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Streptococcus polynucleotide SEQ ID NO 5195.

XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
 XX
 OS Streptococcus pyogenes.

PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB04789.
 XX 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.

XX
 PA (CHIR-) CHIRON SPA.

CC inhibit or activate the activity of the proteins. The antagonists can
 CC be used to treat an individual having need to inhibit a bacterial
 CC protein. Vectors expressing the proteins can be used to induce a
 CC protective immune response in mammals.
 XX
 SQ Sequence 636 BP; 168 A; 144 C; 156 G; 168 T; 0 other;
 Query Match 43.6%; Score 402; DB 18; Length 636;
 Best Local Similarity 98.8%; Pred. No. 8.6e-117; Indels 0; Gaps 0;
 Matches 405; Conservative 0; Mismatches 5;
 QY 1 ATGACTAAACAGCGCTTTTATTTGCTGCTCAAGTGCCAGTATCTAGGATGGGACGG 60
 DB 85 ATGACTAAACAGCGCTTTTATTTGCTGCTCAAGTGCCAGTATCTAGGATGGGACGG 144
 QY 61 GATTTCATGATCAGTATCCGATTTGTCNAAGAAAGCAATGATCGAGCGATCAGTCTC 120
 DB 145 GATTTCATGATCAGTATCCGATTTGTCNAAGAAAGCAATGATCGAGCGATCAGTCTC 204
 QY 121 GGTATGATTTACGTTATCTCATGATACGGAAGAACAACTCAATCAGACCCGCTAT 180
 DB 205 GGTATGATTTACGTTATCTCATGATACGGAAGAACAACTCAATCAGACCCGCTAT 264
 QY 181 ACGCAACAGCAATCTAGCACTCGGTTGCTATCTACCGTTTATTCGAAGAAAGGCG 240
 DB 265 ACGCAACAGCAATCTAGCACTCGGTTGCTATCTACCGTTTATTCGAAGAAAGGCG 324
 QY 241 TATCAGCCTGATATGGTTGCTGTTGCTCTTGAGAAATCTCTGCTTGGTGGCAAGC 300
 DB 325 TATCAGCCTGATATGGTTGCTGTTGCTCTTGAGAAATCTCTGCTTGGTGGCAAGC 384
 QY 301 GCGCCTGATTTTGAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGGAA 360
 DB 385 GCGCCTGATTTTGAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGGAA 444
 QY 361 GAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
 DB 445 GAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
 RESULT 13
 ABA90521
 ID ABA90521 standard; DNA; 2365589 BP.
 XX
 AC ABA90521;
 XX
 DT 16-MAY-2002 (first entry)
 XX
 DE Genomic sequence of *Lactococcus lactis* IL1403.
 XX
 KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese; ds.
 XX
 OS *Lactococcus lactis* IL1403.
 XX
 PN FR2807446-A1.
 XX
 PD 12-OCT-2001.
 XX
 PF 11-APR-2000; 2000FR-0004630.
 XX
 PR 11-APR-2000; 2000FR-0004630.
 XX
 XX (INRG) INRA INST NAT RECH AGRONOMIQUE.
 XX
 PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;
 XX
 DR WPI; 2002-043418/06.
 XX
 PT New nucleotide sequence useful in the identification or *Lactococcus*
 PT *lactis* and related species -
 XX
 PS Claim 1; SEQ ID 1; 2504pp; French.
 XX

CC The present invention is related to a *Lactococcus lactis* nucleotide
 CC sequence (ABA90521) and related proteins (ABA5300-ABA55621). The
 CC nucleic acid sequence is useful in the detection and/or amplification of
 CC nucleic acid sequence, particularly to identify *Lactococcus lactis* or
 CC related species. The proteins of the invention are useful for the
 CC biosynthesis or biodegradation of a composition of interest. The
 CC invention helps research in lactic bacteria, particularly useful in the
 CC production of yogurt and cheese.
 CC Note: The sequence data for this patent is based on equivalent patent
 CC WO200177334 (published 18-OCT-2001) which is available in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
 XX
 SQ Sequence 2365589 BP; 765914 A; 415261 C; 420487 G; 763927 T; 0 other;
 Query Match 42.2%; Score 388.8; DB 24; Length 2365589;
 Best Local Similarity 64.6%; Pred. No. 1.2e-110; Indels 6; Gaps 1;
 Matches 598; Conservative 0; Mismatches 322;
 QY 1 ATGACTAAACAGCGCTTTTATTTGCTGCTCAAGTGCCAGTATCTAGGATGGGACGG 60
 DB 782872 ATGACTAAACAGCGCTTTTATTTGCTGCTCAAGTGCCAGTATCTAGGATGGGACGG 782931
 QY 61 GATTTCATGATCAGTATCCGATTTGTCNAAGAAAGCAATGATCGAGCGATCAGTCTC 120
 DB 782932 GATTTCATGATCAGTATCCGATTTGTCNAAGAAAGCAATGATCGAGCGATCAGTCTC 782991
 QY 121 GGTATGATTTACGTTATCTCATGATACGGAAGAACAACTCAATCAGACCCGCTAT 180
 DB 782992 GGTATGATTTACGTTATCTCATGATACGGAAGAACAACTCAATCAGACCCGCTAT 783051
 QY 181 ACGCAACAGCAATCTAGCACTCGGTTGCTATCTACCGTTTATTCGAAGAAAGGCG 240
 DB 783052 ACTCAACCTGCAATTTTAAACACCTCTGCTGCTTATTTACGTTTGAAGTGAAGTGG 783111
 QY 241 TATCAGCCTGATATGGTTGCTGTTGCTCTCTTGAGAAATCTCTGCTTGGTGGCAAGC 300
 DB 783112 ATTAACCTGACCTTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783171
 QY 301 GCGCCTGATTTTGAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGGAA 360
 DB 783172 GGAATCATTTGATTTTCAAGAACGATTAAGCTTTGAGTAAACGCTGCTCAATATATGACA 783231
 QY 361 GAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 DB 783232 GAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783291
 QY 421 ATGGAAGAGCGCTGCTCAAAAGCTTCTGAACTT-----GGAGTGGTTACTCCAGCCAC 474
 DB 783292 ATGGAAGAGCTTGGCAGAACGCGCGAATTTAAAGGCGGTATTGTGAGTCCAGCAAT 783351
 QY 475 TATAACACACCTGCAAAATCGTCAATGCTGAGAGAGTGGTTGCAAGTATGATCGAGCGGT 534
 DB 783352 TATAACACACCTGCAAAATCGTCAATGCTGAGAGAGTGGTTGCAAGTATGATCGAGCGGT 783411
 QY 535 GAACCTTTTCAAGAACGAGTGGTCCAAACGCTTGAATTCCTTAAAGTGTGAGTCCCTTT 594
 DB 783412 GAGTTGCTTAAAGAGAGCGGAGTTCGTAACCTTATTTGAATTTAAAGTTTTCAGACCTTTC 783471
 QY 595 CACACCGCTCTCTGAGCGCTCTGAGCGAGAACTAGCTGAAACTCTAGCTCAGGTAGT 654
 DB 783472 CATACAGCAATTTTAAACACGATCTGAAATTTGGCTTTTGGAGCTTGATATAATTTGAT 783531
 QY 655 TTTTCAGATTTTACTTGTCCCTAGTCGCGCAATACAGAGCTGCTGATGCAAAAAGAG 714
 DB 783532 TTTTAGACCTTTTGAATTACCATTAATCTCAATACAGTGTCTTAAGTAAATGATATGAT 783591
 QY 715 GACATTGCTCAGCTCTTGACCGCTCAGGTCAAGGAAACCGGTTCTGTTCTATGAAAGTATT 774
 DB 783592 GAAGTCAAAGGACTTTTGACGCGTCAAGTCAAGTCAAGTTCGTTTTTATCAATCGGTT 783651
 QY 775 GGGGTATGCAAGACGAGCATTAACCACTTTATCGAGATTGACCGGGAAGCTCTG 834
 DB 783652 GAAACCAATCAAAAACCTAGGGCGACTCGCTTTTATTGAAGTTGGTCTCGGAGAGTACTT 783711

QY 835 TCAGGTTTCTTAAAAAATGATCAAACTGCTCACTAGCTCATCTGGAAGATCAAGCG 894
DB 783712 TCAGGTTTCTTAAAAAATGATCAAACTGCTCACTAGCTCATCTGGAAGATCAAGCG 783771
QY 895 AGTTTAGTACACTTTTAGAAAAATA 920
DB 783772 TCATTTGAGCTTTGATTATCAGTA 783797

RESULT 14

ABN68640
ID ABN68640 standard; DNA; 924 BP.

XX

AC ABN68640;

XX 01-JUL-2002 (first entry)

DE Streptococcus polynucleotide SEQ ID NO 5193.

KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.

OS Streptococcus agalactiae.

XX WO200234771-A2.

PN 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

XX 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Massignani V, Margarat Ros YI, Grandi G, Fraser C;
XX Tettelin H;

DR WPI; 2002-352536/38.

DR P-PSDB; ABP28009.

PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -

XX Claim 7; Page 3863; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN68640-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.

XX SQ Sequence 924 BP; 311 A; 132 C; 194 G; 287 T; 0 other;

Query Match 42.0%; Score 387; DB 24; Length 924;
Best Local Similarity 64.5%; Pred. No. 6e-112;

Matches 595; Conservative 0; Mismatches 325; Indels 3; Gaps 1;

QY 1 ATGACTAAAAACAGCCCTTTTATTTGCTGTCAGAGTCCCGAGTATCTTAGGATGGGACGG 60

DB 1 ATGAATAAAGTTTCAATTTCTATTGCTGCTCAAGGTGCTCAAAAAGTTTAGGAATGGCTAGA 60

QY 61 GATTTCTATGATCAGTATCCGATGTCAAAAGAAACGATTGATCGAGCGAGTCAGGTGCTC 120

DB 61 GATTTATATGAACTTTCCCGATTGTTTAAAGAAACITTTTATAGGCAAGTCATGTTTTA 120

QY 121 GGTATGATTTACGTTTATCTCATCATACGGAAGAACAAACATCAATCAGACCGCGTAT 180

DB 121 GGATATGATTGAGAGAAATGATTGATAAGAGCTTAGACAAATTAATCAGAGAAATAT 180

QY 181 ACGCAACAGCCATCTTAGCGACTTCGCGTTGCTATCTACCGTTTATTGCG---AAGAAAAG 237

DB 181 ACGCAACAGCTATTTCTAACGACTTCGACTGCTATTATCGTTTAAATTTCTTAAAGAAAT 240

QY 238 GGCATATCAGCGCTGATATGTTGCTGTTGCTCTTGGAGAACTACTCGCTTGGTGCA 297

DB 241 GAATTTAGACAGATATGTTGCGAGACTTCTTTGGAGAAATACAGTGTCTTAGTTGCT 300

QY 298 AGCGGCGCTTGGATTGTTGAAGATGCGGTTGCGCTTGGTAGCTAAGCGTGGAGCCTATG 357

DB 301 TCAGGTGCTATAAGATTGGAAGATGCTGTAGTACTTGTCTAGACGAGGCCAACTAATG 360

QY 358 GAAGAAGCGCTCCTGCTGCTGCTGCAAGATGCTAGCAGTCTCTCAATACGCCAGTAGAG 417

DB 361 GAACAGCAGCACCCTGCTGTTAGTGCGCAAAATGTTAGTAGCAGTTTAAATGCTGATCGTCAA 420

QY 418 GTCATTTGAAGAAGCCTCTCAAAAAGCTTCTGAACTTGGAGTGTACTCCAGCAACTAT 477

DB 421 ATAATTGAAGATGCTTGTGAAGAGGCTTCTCAGTTTGGTATTGTTAGTCTCTGCAAAATTAT 480

QY 478 AACACACTGCACAAATCGTCAITGCTGGAGAGTGGTTCGACTTGTATCGAGCGGTGAA 537

DB 481 AATACTCCAAAGCAAAATGTTTATAGGGGAGAGTCAATCGTCTGTAATGCAAGTGTAGAA 540

QY 538 CTTTGTGAAGAAGCAGGTGCCAAACGCTTGATTCCTCTTAAGGTGTGAGGTCCCTTTTAC 597

DB 541 GAATTTAAACCAACAAGGAGTAAACGTTTAATTCCTCTGAATGTTTCAGGTCTCTTTTAC 600

QY 598 ACCGCTCTCTTGAGCCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGTTAAGTTT 657

DB 601 ACTGCTTTATTTGAAACCTGCAAGTCAGAAATGAGCGATGTTTATAGATAAAGTTCATTTT 660

QY 658 TCAGATTTTACTGTCCCTAGTCCGCAATACAGAACTGCTGTGATGCAAAAAGAGGAC 717

DB 661 TCAGTTTCTGAAATCCCTGTAAATTGGAAATACAGAGGCGCAGATAATGAAAAGACGAT 720

QY 718 ATTGCTCAGCTCTTGACGCGTCAAGTCAAGAACCCGTTCTGTTTCTATGAAAGTATTGGG 777

DB 721 ATCAAAATCTTTATTTGGCACCGCAAGTTATGGAACCTGTAGCCTTTGATGAGTCTATTGAA 780

QY 778 GTCATGCAAGAAGCAGCAGCATTAAGCAACTTTATCGAGATTGGACCGGGGAAAGTCTTCTCA 837

DB 781 ACCATGAAAAAATGGAATGACACAGGTTGTTGAAATTTGTCAGGSAAGGTACTCTCA 840

QY 838 GGTTTGTTAAAAAATTTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGT 897

DB 841 GGTTCCTTAAAAAATATAGATAGTCTCATTCAGTACATTCCTGTGGAAGATAAAATTGGA 900

QY 898 TTAGTAGCCTTTTAGAAAAATA 920

DB 901 TTTAATAATCTTAAGGAATAAA 923

RESULT 15

ABN71527

ID ABN71527 standard; DNA; 2155561 BP.

XX AC ABN71527;

XX XX

02-JUL-2002 (first entry)
Streptococcus polynucleotide SEQ ID NO 10967.
Streptococcus; GAS; group B streptococcus; Streptococcus agalactiae;
group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
Streptococcus sp.
WO200234771-A2.
02-MAY-2002.
29-OCT-2001; 2001WO-GB04789.
27-OCT-2000; 2000GB-0026333.
24-NOV-2000; 2000GB-0028727.
07-MAR-2001; 2001GB-0005640.
(CHIR-) CHIRON SPA.
(GENO-) INST GENOMIC RES.
Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
Tettelin H;
WPI; 2002-352536/38.
New Streptococcus protein for the treatment or prevention of infection
or disease caused by Streptococcus bacteria, such as meningitis, and
for detecting a compound that binds to the protein -
Claim 8; Page 4196-4488; 4525pp; English.
The invention relates to a protein (ABP25413-ABP30895) from group B
Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
(Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
the specification. The proteins have antibacterial and antiinflammatory
activity. (II), nucleic acids encoding (I), ABN6044-ABN71526 and
antibodies that bind (I) are used in the manufacture of medicaments for
the treatment or prevention of infection or disease caused by
Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
Nucleic acids encoding (I) are used to detect Streptococcus in a
biological sample. (I) is used to determine whether a compound binds to
(I). A composition comprising (I) or a nucleic acid encoding (I), may be
used as a vaccine or diagnostic composition. The disease caused by
Streptococcus that is prevented or treated may be meningitis. Nucleic
acid encoding (I) may be used to recombinantly produce (I) and may be
used in gene therapy. Antibodies to (I) are used for affinity
chromatography, immunoassays, and distinguishing/identifying
Streptococcus proteins.
Sequence 2155561 BP; 695741 A; 385320 C; 382958 G; 691542 T; 0 other;
Query Match 42.0%; Score 387; DB 24; Length 2155561;
Best Local Similarity 64.5%; Pred. No. 4.3e-110;
Matches 595; Conservative 0; Mismatches 325; Indels 3; Gaps 1;
QY 1 ATGACTAAACAGCGCTTTTATTTGCTGGTCAAGTCCCGAGTATCTAGGATGGGACGG 60
DB 330947 ATGAATAAGTTTCATTTCTATTTGCTGGTCAAGTGTCAAAAGTTAGGAATGGCTAGA 331006
QY 61 GATTTCATGATCAGTATCCGATTTGTCAGAGAACGATTCGATCAGCGAGTGGTCTC 120
DB 331007 GATTATAGAACTTTCCCGATTTGTTAAAGAACTTTTGAATAGGCAAGTCATGTTTA 331066
QY 121 GGTATGATTACGTTATCTCATCGATACGGAAGACAACTCAATCAGACCCGCTAT 180
DB 331067 GGATATGATTGAGAGAATTCGATTGATAAGACATTAGACAAATTAATCAGAGAAATAT 331126
QY 181 ACGCAACAGCCATCTAGGACATCGGTGCTATCTACCGTTTATTCG---AAGAAAAG 237
DB 331127 ACGCAACAGCTATTCTAAGCACTTCGACTGCTATTATCGTTTAAATTTCTTAAGAAAT 331186

QY 238 GGCTATCAGCCTGATATGTTGCTGTGTTGCTCTTTGGAGAACTACTCTCCTTTGGTGGCA 297
DB 331187 GAACTTAGACCAGATATGTTGCTGAGACTTTCTTTGGGAGAAATACAGTGTCTTTAGTTGCT 331246
QY 298 ACGCGCGCTTGATTTTGAAGATCGGTGCTTTGGTGTAGCTTAAGCGTGAGCCCTATATG 357
DB 331247 TCAGGTGCTATAGATTTGAAGATCTGTAGTACTTTGCTGTAGACGAGCCCACTAAATG 331306
QY 358 GAAGAACGCGCTCCTGCTGACTCTGCAAGATGGTAGCAGTTCTCAATACGCCAGTAGAG 417
DB 331307 GAAGCAGCAGCACCTGCTGCTAGTGGCAAAATGGTAGCAGTTTAAATGCTCATCGTCAA 331366
QY 418 GTCAITGAAGAGCCCTGTCAAAAGCTTCTGAACCTTGAACCTTGAACCTTGAACCTTGAAC 477
DB 331367 ATAATTGAAGATGCTTTGTAAGAGGCTTCTCAGTTTGGTATTTGTTAGTCTCTCAAAATTA 331426
QY 478 AACACACCTGCAACAAATCGTCAATGCTGGAGAGTGGTTGACAGTTGATCGAGCGGTGAA 537
DB 331427 AATACTCCAAAGCAAAATGTTATAGGGGAGAGTCAATCGCTGTAAATGCACTGTAGAA 331486
QY 538 CTTTTCGAAGAGCAGGTGCTCCAAACGCTTGATTCCTCTTAAGGTGTGAGGTCCCTTTTAC 597
DB 331487 GAATTAACCAACAAAGGAGTAAACGTTTAAATCTCTGAATGTTTTCAGGTCCCTTTTAC 331546
QY 598 ACCGCTCTCCTTTGAGCCTGCTAGCCAGAAACTAGCTGAAACTCTAGCTCAGCTAAGTTT 657
DB 331547 ACTGCTTTATTGAACCTGCAAGTCAAGATTTGACGATGTTTATAGATAAAGTTTCAATTT 331606
QY 658 TCAGATTTTACTTTGCTCCCTAGTTCGGCAATACAGAGCTGTGTGATGCAAAAAGAGGAC 717
DB 331607 TCAGTTTCTGAAATCCCTGTAATTTGAAATACAGAGGCGCAGATAATGAAAAAGACGAT 331666
QY 718 ATTGCTCAGCTCTTGACGGGTGAGTCAAGGAACCGTTTCTTCTATGAAAGTATTGGG 777
DB 331667 ATCAATCTTTTATTTGGACGCCCAAGTTATGGAACCTGTACGCTTTGATGAGTCTATTGAA 331726
QY 778 GTCATGCAAGAGCAGGCGATTAAGCAACTTTATCGAGATTGGACCGGGGAAAGCTTTGTCA 837
DB 331727 ACCATGAAAAAATGGGAATGACACAGGTTGTTGAAATTTGGTCCAGGGAAGGTACTCTCA 331786
QY 838 GGTTTTGTAAAAAATTTGATCAAACTGCTACCTTACCTGATGCTGGAAGATCAAGGAGT 897
DB 331787 GGTTCCTTGAAGAAAAATAGATAGCTCATTATCAGTACATTTCTGTTGAAGATAAAATTTGA 331846
QY 898 TTAGTAGCAGCTTTTAGAAAAATA 920
DB 331847 TTTAATAATCTTAAGGAATTHAA 331869

Search completed: June 11, 2003, 18:04:08
Job time : 275 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2003, 17:59:32 ; Search time 70 Seconds
(without alignments)
4034.990 Million cell updates/sec

Title: US-09-308-397-1
Perfect score: 921
Sequence: 1 atcactaaacagcccttttt.....tagcaccttttagaaaaatg 921

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	913	99.1	19702	4	US-08-961-527-7
2	523.2	56.8	1196	4	US-08-858-207A-8
C 3	164.8	17.9	9686	4	US-09-221-017B-1003
4	130.4	14.2	936	4	US-09-134-001C-1987
5	128.4	13.7	927	1	US-08-789-609A-1
6	126.4	13.7	927	1	US-09-108-517-1
7	69.2	7.5	19227	3	US-09-090-793-13
8	69.2	7.5	40138	3	US-09-090-793-12
9	53.4	5.8	8268	1	US-08-375-709-10
10	53.4	5.8	8268	1	US-08-752-929-10
11	53.4	5.8	37895	1	US-08-375-709-1
12	53.4	5.8	37895	1	US-08-752-929-1
13	53.4	5.8	37895	3	US-09-090-793-1
14	53.2	5.8	4403765	4	US-09-103-840A-2
15	53.2	5.8	4411529	4	US-09-103-840A-1
16	47.2	5.1	80161	3	US-09-036-987A-1
17	47.2	5.1	80161	4	US-09-370-700-1
18	45	4.9	345	3	US-09-154-083-9
C 19	40.4	4.4	4403765	4	US-09-103-840A-2
C 20	40.4	4.4	4411529	4	US-09-103-840A-1
C 21	40.2	4.4	2190	4	US-09-625-188-19
22	38.6	4.2	925	3	US-08-858-003-1
23	38.6	4.2	925	3	US-09-078-166-1
24	38.6	4.2	925	4	US-08-997-467-1
25	38.2	4.1	1245	3	US-08-242-050-1
26	38.2	4.1	19702	4	US-08-961-527-7
27	36.6	4.0	15872	4	US-09-105-537-1

28 36.6 4.0 50937 4 US-09-438-517-1 Sequence 1, Appli
29 35.8 3.9 71989 4 US-09-443-501A-2 Sequence 2, Appli
30 35.6 3.9 3279 5 PCT-US93-03077-2 Sequence 2, Appli
31 35.4 3.8 13987 2 US-08-804-227C-13 Sequence 13, Appli
32 35.4 3.8 44377 2 US-08-804-227C-7 Sequence 7, Appli
33 35.4 3.8 44377 2 US-08-804-198-1 Sequence 1, Appli
34 35.2 3.8 13842 4 US-09-105-537-30 Sequence 30, Appli
35 35.2 3.8 38778 4 US-09-105-537-5 Sequence 5, Appli
36 35.2 3.8 38506 3 US-09-320-878-19 Sequence 19, Appli
C 37 34.8 3.8 2824 4 US-07-757-022B-13 Sequence 13, Appli
C 38 34.8 3.8 3066 4 US-07-757-022B-83 Sequence 83, Appli
C 39 34.8 3.8 3117 4 US-07-757-022B-73 Sequence 73, Appli
C 40 34.8 3.8 3148 4 US-07-757-022B-57 Sequence 57, Appli
C 41 34.8 3.8 3420 4 US-07-757-022B-103 Sequence 103, App
C 42 34.8 3.8 3813 4 US-07-757-022B-43 Sequence 43, Appli
C 43 34.8 3.8 3936 4 US-07-757-022B-41 Sequence 41, Appli
C 44 34.8 3.8 3942 4 US-07-757-022B-141 Sequence 141, App
C 45 34.8 3.8 3945 4 US-07-757-022B-49 Sequence 49, Appli

ALIGNMENTS

RESULT 1

US-08-961-527-7/c

; Sequence 7, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340F1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 19702 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-7

Query Match 99.1%; Score 913; DB 4; Length 19702;

Best Local Similarity 99.5%; Pred. No. 8.1e-295;

Matches 916; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ATGACTAAACAGCCCTTTTATTTCTGCTCAAGTGCCCACTATCTAGGATGGGACGG 60

Db 17149 ATGACTAAACAGCCCTTTTATTTCTGCTCAAGTGCCCACTATCTAGGATGGGACGG 17090

Db 431 AAAATTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGCACTT 540
 QY 910 TTAGAAAAATAG 921
 Db 541 TTAGAAAAATAG 552

RESULT 3

US-09-221-017B-1003/c
 ; Sequence 1003, Application US/09221017B
 ; Patent No. 6444799
 ; GENERAL INFORMATION:
 ; APPLICANT: Ross, Bruce C.
 ; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 ; NUMBER OF SEQUENCES: 1120
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 755 PAGE MILL ROAD
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304-1018
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows
 ; SOFTWARE: FastSeq for Windows Version 2.0b
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/221.017B
 ; FILING DATE: 23-DEC-1998
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1182
 ; FILING DATE: 31-DEC-1997
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PP1546
 ; FILING DATE: 30-JAN-1998
 ; APPLICATION DATA:
 ; APPLICATION NUMBER: PP2911
 ; FILING DATE: 09-APR-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/AU98/01023
 ; FILING DATE: 10-DEC-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Monroy, Gladys H
 ; REGISTRATION NUMBER: 32,430
 ; REFERENCE/DOCKET NUMBER: 27340-20021.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-813-5600
 ; TELEFAX: 650-494-0792
 ; TELEX: 706141
 ; INFORMATION FOR SEQ ID NO: 1003:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9686 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: DNA (Genomic)
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: UNKNOWN
 ; ORIGINAL SOURCE:
 ; ORGANISM: PORPHYROMONAS GINGIVALIS
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: 1...9686
 ; US-09-221-017B-1003

Query Match 17.9%; Score 164.8; DB 4; Length 9686;
 Best Local Similarity 52.8%; Pred. No. 2e-44;
 Matches 452; Conservative 0; Mismatches 392; Indels 12; Gaps 4;
 QY 10 ACAGCTTTTATTGCTGGTCAAGGTGCCAGTATCTAGGATGGACGGGATTTCAT 69

Db 4537 AAAGCATTTGTTATCCCGGTTCAGGAGCACAGTTCTGTAGGAATCGGAAAGACCTGTAC 4478
 QY 70 GATCAGTATCCGANTGTCAAAAGAAACGATTGATCGAGCGAGTCTCGTGTATGAT 129
 Db 4477 GAACAGAACCCCGAGGCCAAGCATATTTTGAAGAACCAATGAGATATTAGCTTCGC 4418
 QY 130 TTACGTTATCTCATGATACGAA---GAACAGAACTCAATCAGACCGGTATACGAA 186
 Db 4417 ATTACGGACATCCTTTCAACGGCACGGCAGAGGAGCTGAACAAACAAAGTACCCAG 4358
 QY 187 CCAGCCATTCTAGCACTTCGGTTCCTATCTACCTTTTATTGCAAGAAAGGGCTATCAG 246
 Db 4357 CGGCGATCTTCTCCACTCGTGTATTCTGGCAAGACGATGGGCGAAGA---CTTCCT 4301
 QY 247 CTGATATGTTGCTGGTTTGTCTCTTGAGAACTACTCTGCTGCTGGTGGCAAGCGGCC 306
 Db 4300 CCCGATATGGTGGCAGGCCATTCGCTCGGAGAGTTTCGGCATTTGGTTCGACAGCTGC 4241
 QY 307 TTGGATTTTGAAGATGCGGTTCCTTGGTAGTAAGCGTGGAGCCTATATGAAAGACG 366
 Db 4240 ATGACCTTCGAAGATGCTTCTGCTCTGATCGAAACGTCCTGCCATGCAGAAAGCA 4181
 QY 367 GCTCTGCTGACTCTGGCAAGATGATGAGTTCTCAATACGCCAGTAGAGTCAATTGAA 426
 Db 4180 TCGAAGTACGCCCTCGACTATGCGGCTGTACTGGGTCTGCTGATGAAAGGTGAA 4121
 QY 427 GAAGCTGTCAAAAGCTTCTGAACCTTGGAGTGGTACTCCAGCCCACTATAACACACT 486
 Db 4120 GAAATCTGTGCCCAAGTACCGA---TGAAGTTGCTACCGGCCCACTACACTGCCCG 4064
 QY 487 GCACAAATCGTCATTGCTGGAGAGTGGTTCAGTTGATCGAGCGGTTGAACTTTTGAA 546
 Db 4063 GGACAGATCGTTATATCCGTTCCGTTAGAGAGTGGCAAGAGTGGACAGCTTGGCAACTGCTCAAG 4004
 QY 547 GAAGCAGGTGCCAAGCTTGTATTCCTCTTAAGTGTGCTGCTCCCTTTACACACGCTCTC 606
 Db 4003 GAAGCCGAGCCAGCGTGTCTTAAAGCTCGCTGCTGCTGAGCANTTCCATTCCTCTG 3944
 QY 607 CTTGAGCCTGTAGCCAGAAACTAGCTGAAACTCTAGCTCAGGTAAAGTTTTCAGATTTT 666
 Db 3943 ATGGAACCGGCTCGTGAAGAGTGGCAAAAGCCATCGAGGAGACTACGATCAGTCAGCG 3884
 QY 667 ACTTGTCCCTAGTCGGCAA---TACAGAACTGCTGTATGCAAAAGAGGACATTGCT 723
 Db 3883 ATCTGCCCCCTATATCAAGACGTAAACGGCTTCTGCGCTACCGATCCGCGCAGAGATTAA 3824
 QY 724 CAGCTTTGACGCGTCAGGTCAAGAAACCGTTCGTTTCTATGAAAGTATTGGGTCATG 783
 Db 3823 AAGAACCTGTATGACACAGCTGACGCGACCGGTTCTGCTGACACAGAGTGTCTGAATG 3764
 QY 784 CAAGAGCAGGATTAAGCAACTTTATCGAGATTGACCGGGGAAAGTCTTGTGAGTTT 843
 Db 3763 ACAGCCGATGTTGCCGACCACTTTATGGAGTTAGGCCCGGAAATGTATTGAGGTTTG 3704
 QY 844 GTTAAAAAATGTATC 859
 Db 3703 GTGAAGAAAAATGCTC 3688

RESULT 4

US-09-134-001C-1987
 ; Sequence 1987, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 1987
 ; LENGTH: 936
 ; TYPE: DNA
 ; ORGANISM: Staphylococcus epidermidis
 ; US-09-134-001C-1987

Query Match 14.2%; Score 130.4; DB 4; Length 936;
 Best Local Similarity 50.1%; Pred. No. 1.5e-33;
 Matches 438; Conservative 0; Mismatches 421; Indels 15; Gaps 4;

QY	1	ATGACTTAAACAGCCTTTTATTTGCTGCTCAAGTGCCCGTACTCTAGGATGGGCGG	60
DB	10	ATGGGTAATAAGAGCTTCTCTCCCTGGCCAAAGTGCCACAAAAGTAGGATGGCTAGT	69
QY	61	GATTTCATGATCAGTATCCGATTTGTCAAGAAACGATTGATCGAGCGAGTCAGGTGCTC	120
DB	70	GACTTATATAATGAAGAGAACGCTCGACTCGAGTATTAAATTTAGCTCAAGAGACAGTA	129
QY	121	GGTTATGATTAGGTATCTCATCATAGCAGGAGAGAC---AACTCAATCAGNCCGC	177
DB	130	GACTTTGATTGCTAGAAACGATGTTTACTGATAATGATGGAAATTAGGAGAAACAGAA	189
QY	178	TATACCAACACCGCATTCTAGCGACTTCGGTTGCTATCTACCGTTTATTCGAAGAAAG	237
DB	190	ATAACACACCTGCATTTACACATAGTGTGGTTACTAGATCATTAATCATTTA	249
QY	238	GGGTATCAGCCTGATGTTGCTGCTGTTGCTCTTGGAGATATCTGCTGCTGGTGCA	297
DB	250	AAAGTAGATTACTATGGGACATAGCTTA-----GGAGATATTCAAGCTTAGTTTCT	303
QY	298	AGCGGCGCTTGGATTTCGAAGATCGGTTGCTTGGTAGCTTAACGCTGGAGCCTATATG	357
DB	304	ACTGATGATTATCTTTTGAAGATCGGTTAGAAATGTCGTAAACGTTGGCCAACTATG	363
QY	358	GAAGAAAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	417
DB	364	GCTCAAGCGTTTCTTCAACGCTGCTGAGGATGAGGAGGATGAGGATGAGGATGAGGATG	423
QY	418	GTCATTGAAGAGCGCTGCTCAAAAGCTTCTGAACCTTGGAGTGGTGTACT---CCAGCCAC	474
DB	424	GATGTTGATAGATATGCTCAACGTTTCTCAAAAGACGTTTAAATGAACTGCTTAAT	483
QY	475	TATAACACCTGCAAAATCGTCAATGCTGAGAGAGTGGTTCAGTTCGAGCGGTT	534
DB	484	ATTAACCTACACAGGTCAAATCGTGTGCTGACATATAATCTTTAATTGATGAATTAGTA	543
QY	535	GAATTTTCAAGAGAGCGGTGCGCAACGCTTGTATCTTCTTAAGTGTCAAGTCCCTTT	594
DB	544	GAAAGGCAAGAACTTGGTGTCAACGCTGCTTCCATTAGCTGTTTCCGGTCCCTTT	603
QY	595	CACACGCTCTCCTTGAAGCTGCTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAGT	654
DB	604	CATTCTCAATGATGAAGTATTGAAGAGGATTTTGTCTAATTTCAATTAATTTGAA	663
QY	655	TTTTCAGATTTTACTTGCTCCCTAGTCGCAATACAGAGCTGCTGCTGATGCAAAAGAG	714
DB	664	TGCATAATGCTAAATATATCCAGTGTTCAGATGTTAATGCAAGGAGAGAACCGATGCT	723
QY	715	GACATTGCTCAGCTCTGACGGT---CAGGTCAAGGAAACCGTTTCTTCTATGAAGT	771
DB	724	GAAGTAATTAAACGCAATATGTTAAACAAATTAATTCACCTGTTCAATTTATCAATCA	783
QY	772	ATTGGGGTCAATGCAAGACGAGCATTAAGCAATCTTATCGAGATTGGACCGGGAAAGTC	831
DB	784	ACCGAGTGGTTGATTAATCAAGGTGTGATCATCTTTATTGAAATGGACCGGAAAGTA	843
QY	832	TGTCAGGTTTGTAAAAAATTTGATCAAACTG	865
DB	844	TTATCTGGGCTTATCAAAAAAATAATCGAGATG	877

RESULT 5
 US-08-789-609A-1
 ; Sequence 1, Application US/08789609A
 ; Patent No. 5827689
 ; GENERAL INFORMATION:
 ; APPLICANT: Gentry, Daniel
 ; APPLICANT: Lonsdale, John
 ; APPLICANT: Pearson, Stewart
 ; APPLICANT: Payne, David
 ; TITLE OF INVENTION: No. 5827689el Fabb
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SmithKline Beecham Corporation
 ; STREET: 709 Swedeland Road
 ; CITY: King of Prussia
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19406-0939
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/789,609A
 ; FILING DATE: 24-JAN-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/030685
 ; FILING DATE: 13-NOV-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gimmi, Edward R
 ; REGISTRATION NUMBER: 38,891
 ; REFERENCE/DOCKET NUMBER: GM50004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 610-270-4478
 ; TELEFAX: 610-270-5090
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 927 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; US-08-789-609A-1

Query Match 13.7%; Score 126.4; DB 1; Length 927;
 Best Local Similarity 49.1%; Pred. No. 3.2e-32;
 Matches 45; Conservative 0; Mismatches 456; Indels 15; Gaps 4;

QY	1	ATGACTAAACAGCCTTTTATTTGCTGCTCAAGTGCCCGTACTAGGATGGGACGG	60
DB	1	ATGACTAAACAGCAATTTATTTCCGGGACAAAGTGCCCAAAAGTTGGTATGGCACA	60
QY	61	GATTTCATGATCAGTATCGGATTCGAAGAAACGATTGATCGAGCGAGTCAGGTGCTC	120
DB	61	GATTGTTTAAATAACAATGATCAAGCACTGAATTTTAACTTCAGCAACAAGACGTTA	120
QY	121	GGTTATGATTACGTTATCTCATCGATAGC---GAAGAGACAACTCAATCAGACCCGC	177
DB	121	GACTTGTATTTTAGAGCAATGTTTACTGATGAAGAGGTAATTTGGTGAACCTGAA	180
QY	178	TATACGCAACAGCAATCTTAGGACTTCGGTTGCTATCTACCGTTTATTTGCAAGAAAG	237
DB	181	AACACGCAACAGCTTTATTGACGCATAGTTCGGCATTA-----TTAGCAGCGCTAAA	234
QY	238	GGGTATCAGCTCATATGTTGCTGTTGTTGTTCTCTTGGAGAACTACTCTCCCTTGGTGA	297
DB	235	ATTTTGAATCCTGATTTTACTATGGGGCATAGTTAGGTGAATATTCAGTTAGTTGCA	294
QY	298	AGCGGCGCTTGGATTTGAAGATCGGTTGCTTGGTAGCTTGGTAGCTGAGCGCTATATG	357

Db 295 GCTGACGTATTATCATTTTGAAGATGACAGTTAAATTTCTAGAAAACGTTGTCATTAATG 354
Qy 358 GAAGAGCGGCTCTGCTGACTCTGGCAAGATGCTAGCAGTTCTCAATAGCCAGTAGAG 417
Db 355 GCGCAAGCATTTCTACTGCTGTAGGAGCAGCTGCTGAGTATTTGGGATTTAGATTTT 414
Qy 418 GTCAATTGAAGAGCTGTCA---AAAAGCTTCTGAACTTTGGAGTGGTTACTCCAGGCAAC 474
Db 415 AAAGTCGATGAAATTTTGAAGTCATTATCATCTGATGACAAATAATTTGAACGCAAC 474
Qy 475 TATAACACCTGACAAAATCGTCAATTTCTGCTGGAGAGTGGTTGAGTGTATGAGCGGTT 534
Db 475 ATTAATTTGCCAGGTCAAAATTTGTTTCAGGTGACAAAGCTTTAAATTTGATGAGTAGTA 534
Qy 535 GAACTTTTGAAGAGCAGGTGCCAAAGCTTTGATTTCTCTTAAGGTCAGGTCCCTTT 594
Db 535 GAAAAGGTAATCATTTAGTGCATTTAGTGCATGCTGATGACAAATAATTTGAACGCAAC 594
Qy 595 CACACCGCTCTCTTTGAGCCTGTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAGT 654
Db 595 CATTCATCGCTAATGAAAGTGAATGAAGAGATTTTCAAGTTTACATTAATCAATTTGAA 654
Qy 655 TTTTCAGATTTTACTTCTGCTGCTAGTGGCAATACAGAGCT---GCTGTGATGCAAAA 711
Db 655 TGGCGTAGCTAAGTTCTCTGAGTTCAAAATGTAATGCGCAAGTGAATGACAAA 714
Qy 712 GAGGACATTTGCTAGCTCTTACCGCTGAGTCAAGGACCCGTTCTGTTCTATGAAAGT 771
Db 715 GAAGTAATTAATCTAATATGTTCAAGCAATATATTTACCAGTACATTAATCAATCTCA 774
Qy 772 ATTTGGGTGATGCAAGAGCAGGATAGCAACTTTTACGAGTTGACCGGGGAAAGTC 831
Db 775 ACAGAAATGGCTAATAGACCAAGGTGTTGATCAITTTTATTGAAATTTGCTCTGAAAAGT 834
Qy 832 TTGTGAGCTTTTGTAAAAAATGATCAAACTGCTCACTAGCTCATGTGGAAGATCAA 891
Db 835 TTATCTGGCTAATTAATAAATAAATAGAGTTAAGTTAATCAATTCAACTTTA 894
Qy 892 GCGAGTTAGTACACTTTTAGAAAA 917
Db 895 GAAGATGGAAGATGGAATGAAA 920

RESULT 6

US-09-108-517-1
; Sequence 1, Application US/09108517
; Patent No. 6258934
; GENERAL INFORMATION:
; APPLICANT: Gentry, Daniel
; APPLICANT: Lonsdale, John
; APPLICANT: Pearson, Stewart
; APPLICANT: Payne, David
; TITLE OF INVENTION: No. 6258934el PabD
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,517
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/789,609
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmli, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: GMS0004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 927 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-09-108-517-1

Query Match 13.7%; Score 126.4; DB 4; Length 927;
Best Local Similarity 49.1%; Pred. No. 3.2e-32;
Matches 455; Conservative 0; Mismatches 456; Indels 15; Gaps 4;
Qy 1 ATGACTAAACAGCCCTTTTATTTCTGCTCAAGTGGCCAGTATCTAGGATGGGACGG 60
Db 1 ATGAGTAAACAGCAATTAATTTTCCGGGACAAGTGCCCAAAAGTTGGTATGCAAA 60
Qy 61 GATTTCTATGATCAGTATCCGATTCTCAAGAAACGATTGATCGAGCGAGTCAGGTGCT 120
Db 61 GATTTGTTTAAATAAATGATCAAGCAACTGAAATTTTAACTTCAGCAGCAAGACGTTA 120
Qy 121 GGTATGATTTACGTTATCTCATCGATAG---GAGAGAGACAACTCAATCAGACCCGC 177
Db 121 GACTTTGATATTTTAGAGACAAATGTTTACTGATGAAGAGGTAATTCGGTGAACCTGAA 180
Qy 178 TATAGCGCAACAGCCATTCTAGCGACTTCGGTGTGATCTACCGTTTATTGCAAGAAA 237
Db 181 ACACGCAACAGCCCTTTTATGAGCGATAGTTTCGGCAATTA-----TTAGCAGCGCTAAA 234
Qy 238 GGTATACGCTGATATAGTTGCTGTTGCTCTTGGAGAAATCTCTGCTGTTGGTGCA 297
Db 235 ATTTTGAATCTGATTTTACTATGGGCAATAGTTAGTGGAATATTTCAAGTTTATGTC 294
Qy 298 AGCGGCGCTTGGATTTTGAAGATGCGGTTCCTTGGTAGCTAAGCGTGGAGCCCTATATG 357
Db 295 GCTGACGTATTATCATTTTGAAGATGCGTTAAATTTGTAGAAACGTTGCTCAATTA 354
Qy 358 GAAGAGCGGCTCTGCTGACTCTGGCAAGATGGTAGCAGTTCTCAATACCCAGTAGAG 417
Db 355 GCGCAAGCATTTCTCTACTGCTGAGGAGCATGGCTGCAGTATTTGGGATTTAGATTTT 414
Qy 418 GTCATTGAAGAGCCTGTCA---AAAAGCTTCTGAACTTGGAGTGGTTACTCCAGCCAA 474
Db 415 AAAGTCGATGAAATTTTGAAGTCAATTCATCTGATGACAAATAATTTGAACCCAGCAA 474
Qy 475 TATAACACCTGCAAAATCGTCAATTTCTGAGGAAGTGGTTGAGCTTTGATCGAGCGTT 534
Db 475 ATTAATTTGCCAGGTCAAAATTTGTTTTCAGTTCACAAAGCTTTAAATTTGATGAGCT 534
Qy 535 GAACTTTTGAAGAGCAGGTGCCAAAGCTTGAATTCCTTTAAGGTCGAGGTCCCTTT 594
Db 535 GAAAAGGTAATCATTTAGTGCATAAGCTGTCATGCTTTTAGCAGTATCTGAGCAATTC 594
Qy 595 CACACCGCTCTCTTTGAGCCTGTAGCCAGAACTAGCTGAAACTCTAGCTCAGTAAAGT 654
Db 595 CATTCATCGCTAATGAAGTGAATGAAGAGATTTTCAAGTTTACATTAATCAATTTGAA 654
Qy 655 TTTTCAGATTTTACTTGTCCCTAGTGGGCAATACAGAGCT---GCTGTGATGCAAAA 711
Db 655 TGGCGTAGTGTAAAGTTTCTCTGAGTTCAAAATGTAATGCGCAAGTGAATGACAAA 714
Qy 712 GAGGACATTTGCTAGCTCTTACCGCTGAGTCAAGGACCCGTTCTGTTCTATGAAAGT 771
Db 715 GAAGTAATTAATCTAATATGTTCAAGCAATATATTTACCAGTACATTAATCAATCTCA 774

772 ATTGGGGTCATCGAAGACCGGACATAGCAACTTTATCGAGATTGACCGGGAAAGTC 831
QY
775 ACGAATGGCTAATAGACCAAGGTGTGATCATTTTATGAAATGTGCTCTGGAAGAAGTT 834
Db
832 TTGTCAGGTTTTGGTTAAAAAAAATGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAA 891
QY
835 TTATCTGGCTTAATTAATAAAAAATAAATAGAGATGTTAAGTTAACATCAATTCAAACCTTTA 894
Db
892 GCGAGTTTATAGTACGACTTTTAGAAAA 917
QY
895 GAAGATGTGAAAGGATGGAATGAAAA 920
Db

RESULT 7

```

US-09-090-793-13
; Sequence 13, Application US/09090793
; Patent No. 6140485
; GENERAL INFORMATION:
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
; TITLE OF INVENTION: of polyketide-like synthesis genes in plants
; FILE REFERENCE: CGNE.131.01US
; CURRENT APPLICATION NUMBER: US/09/090,793
; CURRENT FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 19227
; TYPE: DNA
; ORGANISM: Vibrio marinus
US-09-090-793-13

```

Query Match	7.5%; Score 69.2; DB 3; Length 19227;
Best Local Similarity	46.8%; Pred. No. 3.le-12;
Matches 332; Conservative 0; Mismatches 363; Indels 15; Gaps 3;	
QY	152 AAGAGACAACCTCAATCAGACCGCTATACGCAACAGCCAGCCATCTACGACATTCGGTTG 211
Db	
QY	2401 AAGAGAGCAATACGTTTAAAGCAATCGCAACAGCGATGGTAGTTGAGTGTTG 2460
Db	
QY	212 CTATCTACCGTTTATTGCAAGAAAGGGCTATCAGCCGTGATGTTGGTTGCTGGTTTGTCTC 271
Db	
QY	2461 GTCTGTTCAAAACGTTTAAAGCAACAGCGTTTAAAGCTGATTTTGTGTCGGTCATAGTT 2520
Db	
QY	272 TTGAGAAATACTCTGCTTGGTGCAAGCGCGCTTGGATTTTGAAGATGCGGTTGCGCT 331
Db	
QY	2521 TCGGTGATTTAAACGCGATTATGGGCTGCCGATGTTTAAAGCAAGCGATTACATGATGT 2580
Db	
QY	332 TGGTAGCTAAGCGTGGAGCCTATATGAAAGAGCGGCTCTGCTGACTCTGGCAAGATGG 391
Db	
QY	2581 TAGCGGTAGTCGTGGTCAAGCAATGGCTGGCCAGAGCAACAAGATTTTGAT-----G 2634
Db	
QY	392 TAGCAGTTCTCAATACGCCAGTAGAGGTCTTGAAGAAAGCGCTCTCAAAAAGCTTCTGAAC 451
Db	
QY	2635 CAGTAAGATGGCCGCTGTTTGGTGATCCAAAGCAAGTCGCTGTGATCATTTGATACCC 2694
Db	
QY	452 TTGAGTGGTTATCTCCAGCCAACTATAACACACTCGCACAAATGTCATTTGCTGGAGAAG 511
Db	
QY	2695 TTGATGATGCTCTATTGCTAACTTCAACTCGAATAACCAAGTTGTTATTTCGTGGTACTA 2754
Db	
QY	512 TGGTTGCAGTTGATCGAGCGGTGAACCTTTGCAAGAGCAGCGTCCAAACGGTTGATTC 571
Db	
QY	2755 CGGAGCAGGTTGCTGTAGCGGTTACAACCTTAGGTAATGCTGGTTTCAAA---GTTGTC 2811
Db	
QY	572 CTCCTTAAGGTCAGAGTCCCTTTTACACCGCTCTCTCTTGAGCCTGCTAGCCAGAAACCTAG 631
Db	
QY	2812 CACTGCCGGTATCTGCTGGTTCCATACACTTTAGTTGCTCAGCGCAAAACCATTTG 2871
Db	
QY	632 CTGAAACTCTAGCTCAGTAAAGTTTTTCAGATTTTTATCTGTCCTTAGTCGGCAATACAG 691
Db	
QY	2872 CTAAGCGGTTGATAGCGCTAAATTTAAAGCGCCAGCAATTCAGTGTGTTGCTAATGGCA 2931
Db	

Qy	692	AGCTGCTGTG-----ATGCAAAAAGAGGACCAITGCTCAGCTCTTGACGGCGTCAGGTCA	745
Db	2932	CAGCTTGTGTGATTCACGCAACCGAATGACATTTAGAAATAAAGCTGAAAACACCATGC	2981
Qy	746	AGAAACCCGTTCTGTTTCTATGAAAGTATTGGGGTCATGCAAGACGAGCATATAGCAACT	805
Db	2992	TGAAATCTCTTTCAATTCATCAAGAAATTGACAACATCTATGCTGATCGTGGCCGCGTAT	3051
Qy	806	TTATCAGAGATTGACCGGGGAAGTCTTGTGAGGTTTTGTAAAAAAATT	855
Db	3052	TTATCGAATTTGGTCCAAAGATGTATTAACTTAAATCGTTGAAACCAATT	3101

RESULT 8

US-09-090-793-12
; Sequence 12, Application US/09090793
; Patent No. 6140485
; GENERAL INFORMATION:
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
; TITLE OF INVENTION: of polyketide-like synthesis genes in plants
; FILE REFERENCE: CGNE-131.01US
; CURRENT APPLICATION NUMBER: US/09/090,793
; CURRENT FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 40138
; TYPE: DNA
; ORGANISM: Vibrio marinus
; US-09-090-793-12

Query Match.	7.5%;	Score 69.2;	DB 3;	Length 40138;
Best Local Similarity	46.8%;	Pred. No. 5e-12;		
Matches 332;	Conservative	0;	Mismatches 363;	Indels 15; Gaps 3;
Qy	152	AAGAAGACAACTCAATCAGACCCGGTATACGCAACACGACCATTCAGCGACTTCGGTTG	211	
Db	19384	AAGAAGACAACTCAATCAGACCCGGTATACGCAACACGACCATTCAGCGACTTCGGTTG	19443	
Qy	212	CTATCTACCGTTTATTGCAAGAAAAGGGCTATACGCTGTGATATGGTTTGGTGGTTTGTCTC	271	
Db	19444	GTCTGTTCAAAACGTTTAAAGCAAGCAGGTTTTAAAGCTGATTTTCTCCGGTCATAGTT	19503	
Qy	272	TTGCGAAGTACTCTGCTTTGGTGGCAAGCGCGCTTGGATTTTGAAGATGCGGTGCGCT	331	
Db	19504	TCGGTAGTTAACCGCATTTATGGCTGCCGATGTTATGAGCGAAAGCAATTACATGATG	19563	
Qy	332	TGCTAGCTAAGCGTGGAGCCTATATGGAAGAACGGCTCTCTGCTGACTCTTGGCAAGATGG	391	
Db	19564	TAGCGCGTAGTCGTGGTCAAGCAATGGCTCGCCAGAGCAACAGATTTTGAT-----G	19617	
Qy	392	TAGCAGTTCTCAATAGCCAGCTAGAGGTCAATTGAAGAGCCTGTGCAAAAGCTTCTGAAC	451	
Db	19618	CAGGTAAGATGGCGCGTGTGTTGGTGATCTCAAGCAAGTCGCTGTGATCATTTGATACCC	19677	
Qy	452	TTGGAGTGGTTACTCCAGCGCAACTATAACACACCTTGCAAAATCGTCTATTGCTGGAGAAG	511	
Db	19678	TTGATGATGTCCTATTGTTACTTCACTCGAATAA CCAAGTTGTTATTGCTGGTACTA	19737	
Qy	512	TGGTTGAGTTGATCAGAGCGGTTGAACCTTTTGCAGAAAGCAGGTGCGCAACCGTTGATTC	571	
Db	19738	CGGAGCAGGTTGCTGTAGCGGTTACAACTTAGGTAATGCTGGTTTCAAA---GTTGTGC	19794	
Qy	572	CTCTTAAGGTGCAGGTGCTCTTTACACCGCTCTCCTTTGAGCGCTGCTAGCCAGAACTAG	631	
Db	19795	CATGCGGTAATCTGTGGTTCCATACACCTTTAGTTTCGTACGCGCAAAACCATTTG	19854	
Qy	632	CTGAACCTCTAGCTCAGGTAAGTTTTTCAGATTTTACTTGTCCCTTAGTTCGGCAATACAG	691	

Db 19855 CTAAAGCGGTGATAGCGCTAAATTTAAAGCGCAAGCAATTCAGTGTGTTGCTAATGGCA 19914
 QY 692 AAGTGTGTG-----ATGCAAAAGAGGACATTTGCTAGCTTGAACCGGTGAGTCA 745
 Db 19915 CAGCTTGGTGCATTCAGCAACCGAATGACATTAAGAAAAACCTGAAAAACCATGC 19974
 QY 746 AGGAACCGTTCGTTTCTATGAAAGTATTGGGGTCATGCAAGAGGAGGCAATGAAGCAACT 805
 Db 19975 TGGAACTGTTCATTTCAATCAAGAAATTGACACATCTATGCTGATGGTGGCGGGTAT 20034
 QY 806 TTATCGAGATGGACCGGGGAAATCTTGTGAGGTTTGTGTAATAAAATTT 855
 Db 20035 TTATCGAATTTGGTCCAAAGAAATGATTAATACTAAATTTGGTTGAAACATTT 20084

RESULT 9

US-08-375-709-10
 ; Sequence 10, Application US/08375709
 ; Patent No. 5683898
 ; GENERAL INFORMATION:
 ; APPLICANT: YAZAWA, Kazunaga
 ; APPLICANT: YAMADA, Akiko
 ; APPLICANT: KATO, Seishi
 ; APPLICANT: KONDO, Kiyosi
 ; TITLE OF INVENTION: Gene Coding For Eicosapentaenoic Acid
 ; TITLE OF INVENTION: Synthesizing Enzymes and Process for Production of
 ; TITLE OF INVENTION: Eicosapentaenoic Acid
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/375,709
 ; FILING DATE: 20-JAN-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/178,251
 ; FILING DATE: 14-MAY-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 4-147945
 ; FILING DATE: 15-MAY-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: WEGNER, Harold C.
 ; REGISTRATION NUMBER: 25,258
 ; REFERENCE/DOCKET NUMBER: 53466/150/AACK
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)672-5300
 ; TELEFAX: (202)672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8268 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ORIGINAL SOURCE:
 ; ORGANISM: Shewanella putrefaciens SCRC-2874 (FERM
 ; ORGANISM: BP-1625)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..8268
 ; FEATURE:
 ; NAME/KEY: mat_peptide

; LOCATION: 1..8268
 US-08-375-709-10
 Query Match 5.8%; Score 53.4; DB 1; Length 8268;
 Best Local Similarity 46.7%; Pred. No. 3.6e-07;
 Matches 260; Conservative 0; Mismatches 276; Indels 21; Gaps 2;
 QY 152 AAGAGAGCAAACTCAATCAGACCCGCTATACGCAACCAAGCCATCTACGCACTTCGGTTG 211
 Db 1991 AAGAAGCCATTTTGACCAATACCGCAATGCGCAAGCGCAATGCTGCGATTCGAATGG 2050
 QY 212 CTATCTACCGTTTATTGCAAGAAAAGGGCTATACGCTGTATATGGTTCTGCTGTTGCTC 271
 Db 2051 GTCATACGATTTGTTTACTGCGGCTGCTTTAATGCGACATGTTGCGAGCCATAGCT 2110
 QY 272 TTGGAGAATACTCTGCTTGGTGGCAAGCGCGCTTGGATTTGAAGATGGGTTGCT 331
 Db 2111 TTGGTGAGCTAAGTGCATCTGTGCTGCAGGTGTTTATTTTCAGCTGATGACTACTACAAG 2170
 QY 332 TGGTAGCTAAGCGTGGAGCCTATATGGAAGAAAGCGGCTCCCTGCT----- 375
 Db 2171 TGGCTTTTGGCTGCTGAGGCTATGCGCAACAAAAGCACCGCTAAAGACGCGTTGAAG 2230
 QY 376 --GACTCTGGCAAGATGTGTAGAGTTCTCAATACCCAGTAGAGTCAATTGAAGAGCT 433
 Db 2231 CAGATGCGAGGAGCAATGTTTGCATCATAACCAAGAGTGTGCGAGACCTTGAAACCGTTG 2290
 QY 434 GTCAAAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGGCCAACTATAACACACCTGCACAAA 493
 Db 2291 AAGCCACCATCGCTAAATTTGATGGGTGAAAGTGCCTAATTAACCGCGCAACGCAAT 2350
 QY 494 TCGTCATTGCTGGAGAAGTGGTTGCAAGTGTGCGAGCGGTTGAACTTTTGAAGAAAGCAG 553
 Db 2351 CAGTAATTGCGAGGCCAACAGCAACTACCGCTGATGCGGCTAAAGCGCTAACTGAGCTTG 2410
 QY 554 GTGCCAAACGCTTGATCTCTTAAGGTGTCAGGTCCCTTTACACCCGCTCTCCCTTGAGC 613
 Db 2411 GTTACAAA---GCGATTAACTGCCAGTATCAGGTGCATTCACACTGAATTTGTTGTC 2467
 QY 614 CTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAAGTTTTTCAGATTTTACTTCTC 673
 Db 2468 ACGCTCAAGCGCCATTTGCTAAAGCGATTGAGCGAGCCAAATTTACTAAACCAAGCCGAG 2527
 QY 674 CCTAGTGGCAATACA 690
 Db 2528 CACTTTACTCAATGCA 2544
 RESULT 10
 US-08-752-929-10
 ; Sequence 10, Application US/08752929
 ; Patent No. 5798259
 ; GENERAL INFORMATION:
 ; APPLICANT: YAZAWA, Kazunaga
 ; APPLICANT: YAMADA, Akiko
 ; APPLICANT: KATO, Seishi
 ; APPLICANT: KONDO, Kiyosi
 ; TITLE OF INVENTION: Gene Coding for Eicosapentaenoic Acid Synthesizing
 ; TITLE OF INVENTION: Enzymes and Process for Production of Eicosapentaenoic
 ; TITLE OF INVENTION: Acid
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/752,929
;; FILING DATE: 20-NOV-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/375,709
;; FILING DATE: 20-JAN-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/178,251
;; FILING DATE: 14-MAY-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 4-147945
;; FILING DATE: 15-MAY-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: WEGNER, Harold C.
;; REGISTRATION NUMBER: 25,258
;; REFERENCE/DOCKET NUMBER: 53466/150/AAOK
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEFAX: (202)672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8268 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Shewanella putrefaciens SCRC-2874 (FERM
;; ORGANISM: BP-1625)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 1..8268
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 1..8268
;; US-08-752-929-10

Query Match 5.8%; Score 53.4; DB 1; Length 8268;
Best Local Similarity 46.7%; Pred. No. 3.6e-07;
Matches 260; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 152 AAGAAGCAAACTCAATCAGACCCGCTATACGCAACCGCATTTCTAGGACCTTGGTTG 211
DB 1991 AAGAAGCCATTTTGACCAATACCGCAATGCCAAAGCCCAATGGTGGCATTTCAATGG 2050
QY 212 CTATCTACCGTTTATTGCAAGAAAGGGCTATCAGCCTGATATGTTGTTGTTTGTCTC 271
DB 2051 GTCAATACGATTGTTTACTGCGGCTGGCTTTAATGCCACATGTTGCGAGCCATAGCT 2110
QY 272 TTGGAGAACTACTCGCTTGGTGGCAAGCGGCGCTTGAATTTTGAAGATGCGGTTGCCT 331
DB 2111 TTGGTGAGCTAAGTGACCTGTGTGCTGCAAGTGTATTTCAGCTGATGACTACTACAAGC 2170
QY 332 TGGTAGTAAGCGTGAGCGCTATATGCAAGAGCGCTCTCTCT----- 375
DB 2171 TGCTTTTCTGCTGGTGGCTATGSCCAAAAGCAGCCGCTAAAGCGCGTTGAG 2230
QY 376 --GACTCTGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAGGTATTGAAGAGCGCT 433
DB 2231 CAGATCAGAGCAATGTTTGAATCATACCAAGAGTCTGCAGACCTTGAACCGTTG 2290
QY 434 GTCAAAAGCTTCTGAAGTTGGATGGTGTACTCCAGCCCACTATACACACTGCGACAAA 493
DB 2291 AGCCACCATCGCTAAATTTGATGGGTGAAGTGTCTTAATACGCGCCCAACGCAAT 2350
QY 494 TCGTCAATGCTGGAGAAGTGGTTGCAAGTGTATCGAGCGGTTGAACTTTTGAAGAGAGAG 553
DB 2351 CAGTAATTGAGGCCCAACAGCAACTACCGCTGATGCGGCTAAAGCGCTAACTAGAGTTG 2410
QY 554 GTGCCAAAGCTGTGATTCCTCTTAAGTGTGAGGTCCTTTTACACCGCTCTCTCTGAGC 613
DB 152 AAGAAGCAAACTCAATCAGACCCGCTATACGCAACCGCATTTCTAGGACCTTGGTTG 211
DB 15896 AAGAAGCCATTTTGACCAATACCGCCCAATGCGCAATTTGGTGGCAATTTCAATGG 15955

Db 2411 GTTACAAA--GCGATTAACTGCCAGTATCAGGTGCAATCCACACTGAACCTTGGTGC 2467
QY 614 CTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAGTTTTCAGATTTTACTTGTG 673
Db 2468 ACGCTCAAGCGCCATTTGCTAAAGCGATTGACGAGCCAAATTTACTAAAACAAGCCGAG 2527
QY 674 CCCTAGTCGGCAATACA 690
Db 2528 CACTTTACTCAATGCA 2544

RESULT 11
US-08-375-709-1
; Sequence 1, Application US/08375709
; Patent No. 5683898
; GENERAL INFORMATION:
; APPLICANT: YAZAWA, Kazunaga
; APPLICANT: YAMADA, Akiko
; APPLICANT: KATO, Seishi
; APPLICANT: KONDO, Kiyosi
; TITLE OF INVENTION: Gene Coding For Bicosapentaenoic Acid
; TITLE OF INVENTION: Synthesizing Enzymes and Process for Production of
; TITLE OF INVENTION: Bicosapentaenoic Acid
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,709
; FILING DATE: 20-JAN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,251
; FILING DATE: 14-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-147945
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/150/AAOK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37895 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Shewanella putrefaciens SCRC-2874 (FERM
; ORGANISM: BP-1625)
; US-08-375-709-1

Query Match 5.8%; Score 53.4; DB 1; Length 37895;
Best Local Similarity 46.7%; Pred. No. 9.3e-07;
Matches 260; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 152 AAGAAGCAAACTCAATCAGACCCGCTATACGCAACCGCATTTCTAGGACCTTGGTTG 211
DB 15896 AAGAAGCCATTTTGACCAATACCGCCCAATGCGCAATTTGGTGGCAATTTCAATGG 15955

QY 212 CTATCTACCGTTTAAATGCAAGAAAGGCGCTATACGCTGATATATGTTGCTGTTGCTC 271
Db 15956 GTCAATACGATTTGTTTACTCGGCTGCTTAAATGCGCATGTTGCGAGCGCATAGCT 16015
QY 272 TTGGAGATATCTCTGCTTGTGTCGAGCGCGCTTGGATTTTGAAGATCGGTTGCT 331
Db 16016 TTGGTACCTAAGTCACTGTGTGTCGAGTGTATTTTACGTGATGACTACTACAAGC 16075
QY 332 TGGTAGCTAAAGCGGAGCTTATATGGAAGAGCGGCTCTGCT----- 375
Db 16076 TGGCTTTTGTCTGCTGAGGCTATGCGACAAAGACCGCTAAAGACGGCGTTGAAG 16135
QY 376 --GACTCTGGCAAGTGTAGCTTCTCAATACCGAGTAGAGTCAATGAAGACCT 433
Db 16136 CAGTGCAGGAGCAATGTTTCAATCATACCAAGAGTGTGCGAGCTTGAACCGTTG 16195
QY 434 CTCAAAAAGCTTCTGAATCTGGAGTGTGTTACTCCAGCCAACTATAACACACCTGCACAAA 493
Db 16196 AAGCCACCATCGTAAATTTGATGGGTGAAGTGCCTAACTATAACGGCCACGCAAT 16255
QY 494 TCGTCATTGCTGGAGAGTGTGTCAGTGTGATCGAGCGGTTGAACCTTTGCAAGAACGAG 553
Db 16256 CAGTAATTGCGAGGCCAACACAGCAACTACCGCTGATGCGGCTAAAGCGCTAACTGAGCTTG 16315
QY 554 GTGCCAAAGCTTGATTCCTCTTAAGGTGTGAGGTGTCAGGTCCTTTTACACACCGCTCTCCTTGAGC 613
Db 16316 GTTACAAA--CGGATTAACTGCCAGTATCAGGTGCAATCCACACTGAACTTTGGTC 16372
QY 614 CTGTAGCCAGAACTAGCTGAACTCTAGCTCAGTAAAGTGTTCAGATTTTACTTTGTC 673
Db 16373 AGCTCAAGCGCAATTTGCTAAAGCGATTGACGCGCAAAATTTACTAAAAACAGCGGAG 16432
QY 674 CCTAGTGGCAATACA 690
Db 16433 CACTTTACTCAATGCA 16449

RESULT 12
US-08-752-929-1
; Sequence 1, Application US/08752929
; Patent No. 5798259
; GENERAL INFORMATION:
; APPLICANT: YAZAWA, Kazunaga
; APPLICANT: YAMADA, Akiko
; APPLICANT: KATO, Seishi
; APPLICANT: KONDO, Kiyosi
; TITLE OF INVENTION: Gene Coding for Eicosapentaenoic Acid Synthesizing
; TITLE OF INVENTION: Enzymes and Process for Production of Eicosapentaenoic
; TITLE OF INVENTION: Acid
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/375,709
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/178,251
; FILING DATE: 14-MAY-1993

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-147945
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 53466/150/AAOK
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 37895 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Shewanella putrefaciens SCRC-2874 (FERM
; ORGANISM: BP-1625)
; US-08-752-929-1
Query Match 5.8%; Score 53.4; DB 1; Length 37895;
Best Local Similarity 46.7%; Pred. No. 9.3e-07;
Matches 260; Conservative 0; Mismatches 276; Indels 21; Gaps 2;
QY 152 AAGAGACAACTCAATCAGACCCGCTATACGACCAACGCGCTTCTAGCGACTTCGGTTG 211
Db 15896 AAGAGACCAATTTGACCAATACCGCAATGCCAAAGCCCAATGGTGGATTTCAATGG 15955
QY 212 CTATCTACCGTTTAAATGCAAGAAAGGCTATACGCTGATATGTTGCTGTTGCTC 271
Db 15956 GTCAATACGATTTGTTTACTGCGGCTGCTTAAATGCGACATGTTGCGAGCCATAGCT 16015
QY 272 TTGGAGATATCTCTGCTTGTGTCGAGCGCGCTTGGATTTTGAAGATCGGTTGCT 331
Db 16016 TTGGTACCTAAGTCACTGTGTGTCGAGTGTATTTTACGTGATGACTACTACAAGC 16075
QY 332 TGGTAGCTAAAGCGGAGCTTATATGGAAGAGCGGCTCTGCT----- 375
Db 16076 TGGCTTTTGTCTGCTGAGGCTATGCGCAAAAGACCGCTAAAGACGGCGTTGAAG 16135
QY 376 --GACTCTGGCAAGTGTAGCTTCTCAATACCGAGTAGAGTCAATGAAGACCT 433
Db 16136 CAGTGCAGGAGCAATGTTTGAATCATACCAAGAGTGTGCGAGCTTGAACCGTTG 16195
QY 434 CTCAAAAAGCTTCTGAATCTGGAGTGTGTTACTCCAGCCAACTATAACACACCTGCACAAA 493
Db 16196 AAGCCACCATCGTAAATTTGATGGGTGAAGTGCCTAACTATAACGGCCACGCAAT 16255
QY 494 TCGTCATTGCTGGAGAGTGTGTCAGTGTGATCGAGCGGTTGAACCTTTGCAAGAACGAG 553
Db 16256 CAGTAATTGCGAGGCCAACACAGCAACTACCGCTGATGCGGCTAAAGCGCTAACTGAGCTTG 16315
QY 554 GTGCCAAAGCTTGATTCCTCTTAAGGTGTGAGGTGTCAGGTCCTTTTACACACCGCTCTCCTTGAGC 613
Db 16316 GTTACAAA--CGGATTAACTGCCAGTATCAGGTGCAATCCACACTGAACTTTGGTC 16372
QY 614 CTGTAGCCAGAACTAGCTGAACTCTAGCTCAGTAAAGTGTTCAGATTTTACTTTGTC 673
Db 16373 AGCTCAAGCGCAATTTGCTAAAGCGATTGACGCGCAAAATTTACTAAAAACAGCGGAG 16432
QY 674 CCTAGTGGCAATACA 690
Db 16433 CACTTTACTCAATGCA 16449

RESULT 13
US-09-090-793-1
; Sequence 1, Application US/09090793
; Patent No. 6140486
; GENERAL INFORMATION:

```
; APPLICANT: Calgene, LLC
; TITLE OF INVENTION: Production of polyunsaturated fatty acids by expression
; FILE REFERENCE: CGNE.131.01US
; CURRENT APPLICATION NUMBER: US/09/090,793
; CURRENT FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,650
; EARLIER FILING DATE: 1997-06-04
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 37895
; TYPE: DNA
; ORGANISM: Shewanella putrefaciens
US-09-090-793-1

Query Match      5.8%; Score 53.4; DB 3; Length 37895;
Best Local Similarity 46.7%; Pred. No. 9.3e-07;
Matches 260; Conservative 0; Mismatches 276; Indels 21; Gaps 2;

QY 152 AGAAGACAACTCAATCAGACCCGCTATACGCAACCGCCATTCAGGACTTCGGTTG 211
Db 15896 AGAAGCCATTTTGACCAATACCGCAATGCCAAAGCGCAATTCGTCGATTTCAATGG 15955
QY 212 CTATCTACCGTTTATTGCAAGAAAGGGCTATCAGCCTGATATGTTGCTGTTTGTCTC 271
Db 15956 GTCAATACGATTGTTTACTGCGGCTGGCTTTAATGCCGACATGTTGAGGCCATAGCT 16015
QY 272 TTGAGAAATACTCTGCTTGGTGGCAGCGGCCCTTGGATTTTGAAGATCGGTTGGCT 331
Db 16016 TTGGTGAGCTAAGTGACATGTGTGTCGAGGTGTTATTTACGCTGATGACTACTACAAG 16075
QY 332 TGCTAGCTAAGCGTGAGCGCTATATGGAAGAGCGGCTCTCTCT----- 375
Db 16076 TGCTTTTCTGCTGTGAGGCTATGGCAACAAAGCAGCCGCTAAAGCGCGTTGAAG 16135
QY 376 --GACTCTGCGAAGATGTAGCAGTTCTCAATACGCCAGTAGAGTCAATTGAAGAGCCT 433
Db 16136 CAGATCAGGAGCAATGTTTGAATCATACCAAGAGTCTGCAGACCTTGAAACCGTTG 16195
QY 434 GTCAAAAGCTTCGAATCGGATGGTGTACTCCAGCCAACTATAACACACCTGCACAAA 493
Db 16196 AAGCCACCTCGCTAAATTTGATGGGTGAAGTGTCTTAATTAACGGCCCAACGCAAT 16255
QY 494 TCCTCATTCGTGAGAGTGGTTGTCAGTTGATCGAGCGGTGAACTTTTGAAGAGCAG 553
Db 16256 CAGTAATTCAGGCCCAACAGCAACTACCGCTGATCGGCTAAAGCGTAACTGAGCTG 16315
QY 554 GTCCCAACGCTTGATTCCTCTTAAGGTGTGAGTCCCTTTTCCACCCGCTCTCCTTGAGC 613
Db 16316 GTTACAAA--GCGATTAACTGCCAGTATCAGGTGCAATCCACACTGAACCTGTTGTGC 16372
QY 614 CTGCTAGCCAGAACTAGCTGAACTCTAGCTCAGGTAGTATTTTTCAGATTTTACTTGT 673
Db 16373 ACCTCAAGCGCCATTGCTAAAGCGATTGACGCGCAAAATTTACTAAACACGCCAG 16432
QY 674 CCTAGTCGCAATACA 690
Db 16433 CACTTTACTCAATGCA 16449

RESULT 14
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      5.8%; Score 53.2; DB 4; Length 4411529;
Best Local Similarity 49.4%; Pred. No. 2.2e-05;
Matches 182; Conservative 0; Mismatches 173; Indels 15; Gaps 1;
```

```
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      5.8%; Score 53.2; DB 4; Length 4403765;
Best Local Similarity 49.4%; Pred. No. 2.2e-05;
Matches 182; Conservative 0; Mismatches 173; Indels 15; Gaps 1;

QY 7 AAAACAGCCCTTTTATTGCTGTCAAGGTGCCAGTATCTAGGGATGGACGGGATTTC 66
Db 489058 AAGACGGCTTTTGTGTTTCCGCTCAGGGCTCGCACTGGGTATGGGTAGCAGCTT 489117
QY 67 TATGATCAGTATCCGATGTGCAAAAGACGATTGATCGAGCGAGTCAGGTGCTCGGTTAT 126
Db 489118 TATGCTGCCTACCCGGTTTTCGCCGAGGCCCTCGATGCTGTGTTGACGAGTTGACCGG 489177
QY 127 GATTTACGTTATC-----TCATCGATACGGAAGAAAGACAACTCAATCAG 171
Db 489178 CACTGCGGTATCCGCTGCGGATGATCTGGGGCAGACCAAGATCTGTTGAATACC 489237
QY 172 ACCCGCTATACGAAACAGCCATTCAGCACTTCGGTGTGCTATCTACCGTTTATGCAA 231
Db 489238 ACCGAATTCGCCAGCGCGCTGTTTGGGTGAGGTGGCGCTGATCGGCTCATG 489297
QY 232 GAAAAGGCTATCAGCTGATATGTTGCTGTTTCTCTTGAGAACTCTGCTTG 291
Db 489298 TCGTGGGGGTGCGCGCGGTTTGGTCTGGGTTCATTCGTTGGCGAGTTGCCCGCGG 489357
QY 292 GTGCAAGCGCGCCCTTGGATTTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGAAGCC 351
Db 489358 CACGTGCGCGGGCGCTGTGTTTCCCGGATGCGCGCATGCTGTGCGCGCGTGACGG 489417
QY 352 TATATGGAAG 361
Db 489418 TTGATGAGG 489427

RESULT 15
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      5.8%; Score 53.2; DB 4; Length 4411529;
Best Local Similarity 49.4%; Pred. No. 2.2e-05;
Matches 182; Conservative 0; Mismatches 173; Indels 15; Gaps 1;
```

```
Qy 7 AAAAAAGCCTTTTATTTGCTGCTCAGGTGCCCGAGTATCTAGGATGGACGGATTTC 66
Db 487616 AAGACGGCTTTTGTTGCTGCTCAGGTGCCCGAGTATCTAGGATGGACGGATTTC 487675
Qy 67 TATGATCAGTATCCGATTTGCAAGAAACGATTGATCCAGCGAGTCAAGTGTGCTCGGTTAT 126
Db 487676 TATGCTGCTTACCCGGTTTTCGCCGAGGCCCTCGATCTGTGTGGACGAGTTGGACCG 487735
Qy 127 GATTACGTTATC-----TCATCGATAGGGAAGAACAACTCAATCAG 171
Db 487736 CACCTGGGGTATCCGCTGCGGATGTGATCTGGGGGACGACCAAGATCTGTTGAATACC 487795
Qy 172 ACCCGCTATAGCGCAACAGCCATTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAA 231
Db 487796 ACCGNATTCGCCAGCCGCGCTGTTGCGGTGAGGTGGCGTGATCCGCTGCTCATG 487855
Qy 232 GAAAAGGCTATCAGCCTGATATCGTTGCTGTTGTTGCTCTTTGGAGAAATACTCTGCCCTG 291
Db 487856 TCGTGGGGGTGCGCGGGGTTTGGTGTGGGTCAATTCGGTGGCGAGTTGGCGCGCG 487915
Qy 292 GTGCAAGCGCGCCTTGGATTTTGAAGATCGCGTTGCCCTTGGTAGCTAAGCGTGGAGCC 351
Db 487916 CACGTCCCGGGCGCTGTGTTGCCGATGCGCGGATGCTGTGGTGGCGCGGTGGACGG 487975
Qy 352 TATATGGAAG 361
Db 487976 TTGATGCAGG 487985
```

Search completed: June 11, 2003, 19:47:06
Job time : 88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: June 11, 2003, 18:44:57 ; Search time 165 Seconds
(without alignments)
7811.963 Million cell updates/sec

Title: US-09-308-397-1
Perfect score: 921
Sequence: 1 atgacataaacagccctttt.....tagcacttttagaaaaaatg 921

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 69976893 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	911.4	99.0	921	10	US-09-815-242-9482
2	344.4	37.4	942	10	US-09-815-242-6823
3	340.4	37.0	3656	10	US-09-070-927A-577
4	204	22.1	950	10	US-09-974-300-1143
5	204	22.1	960	10	US-09-974-300-1086
6	196.4	21.3	939	10	US-09-815-242-6878
7	196.4	21.3	1830121	9	US-10-329-960-1
8	185.6	20.2	930	10	US-09-815-242-6028
9	158	17.2	6021	10	US-09-070-927A-458
10	139.6	15.2	592	10	US-09-974-300-5552
11	128	13.9	936	10	US-09-815-242-8032
12	127.6	13.9	906	10	US-09-815-242-4405
13	120.8	13.1	939	10	US-09-815-242-7796
14	118.4	12.9	6251	7	US-08-781-986A-25
15	69.2	7.5	7959	9	US-10-331-061-77
16	69.2	7.5	19279	9	US-10-331-061-13
17	89.2	7.5	40138	9	US-10-331-061-12
18	53.4	5.8	8268	9	US-10-331-061-83
19	53.4	5.8	37895	9	US-10-331-061-1

20	53.2	5.8	4209	9	US-09-712-363-20
21	51.8	5.6	64681	10	US-09-790-388-1
22	47	5.1	930	10	US-09-815-242-7193
23	40	4.3	268	10	US-09-815-242-1745
24	38.6	4.2	925	9	US-09-735-056-1
25	38.4	4.2	1236	9	US-10-077-584-3
26	38.2	4.1	1380	9	US-10-074-045-24
27	37.4	4.1	4851	9	US-09-712-363-116
28	37.2	4.0	532	9	US-10-184-644-564
29	37.2	4.0	532	9	US-10-184-644-564
30	36.6	4.0	520	9	US-10-184-644-332
31	36.6	4.0	520	9	US-10-184-644-332
32	36.6	4.0	15872	9	US-09-860-846-1
33	36.6	4.0	15872	9	US-09-988-384B-1
34	36.6	4.0	15872	9	US-09-836-821-1
35	36.6	4.0	15872	10	US-09-861-289-1
36	36.6	4.0	50937	9	US-09-808-880-1
37	35.6	3.9	3282	10	US-09-864-864-298
38	35.4	3.8	5558	10	US-09-070-927A-11
39	35.2	3.8	13842	9	US-09-860-846-30
40	35.2	3.8	13842	9	US-09-988-384B-30
41	35.2	3.8	13842	9	US-09-836-821-30
42	35.2	3.8	13842	10	US-09-861-289-30
43	35.2	3.8	36778	9	US-09-860-846-5
44	35.2	3.8	36778	9	US-09-836-821-5
45	35.2	3.8	36778	10	US-09-861-289-5

ALIGNMENTS

RESULT 1
US-09-815-242-9482
; Sequence 9482, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Chisen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9482
; LENGTH: 921
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(921)
US-09-815-242-9482

Query Match 99.0%; Score 911.4; DB 10; Length 921;
Best Local Similarity 99.3%; Pred. No. 8.9e-280;
Matches 915; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ATGACTAAAACAGCCTTTTATTTGCTGTCAGGTGCGGATCTAGGATCGGACGG 60
Db 1 ATGACTAAAACAGCCTTTTATTTGCTGTCAGGTGCGGATCTAGGATCGGACGG 60

QY 61 GATTTCATGATCAGTATCCGATGTCAAAGAAACGATTGATCGAGCGAGTCAGTGTCTC 120
Db 61 GATTTCATGATCAGTATCCGATGTCAAAGAAACGATTGATCGAGCGAGTCAGTGTCTC 120

QY 121 GGTATGATTACGTTATCTCATGATCGGAGAGACAACTCAATCAGACCGCTAT 180
Db 121 GGTATGATTACGTTATCTCATGATCGGAGAGAGACAACTCAATCAGACCGCTAT 180

QY 181 ACGCAACAGCCATTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240
Db 181 ACGCAACAGCCATTCTAGCGACTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGCG 240

QY 241 TATCAGCCTGATATGGTTGCTGGTTGCTCTTGGAGATATCTGCGCTTGGTGGCAGC 300
Db 241 TATCAGCCTGATATGGTTGCTGGTTGCTCTTGGAGATATCTGCGCTTGGTGGCAGC 300

QY 301 GCGCGCTTGGATTTGAAGATCGGTTGCTTGGTAGCTTAAGCGTGGAGCCTATATGAA 360
Db 301 GCGCGCTTGGATTTGAAGATCGGTTGCTTGGTAGCTTAAGCGTGGAGCCTATATGAA 360

QY 361 GAAGCGGCTCTGCTGACTCTGGAGAGATGGTAGAGTTCTCAATACGCCAGTAGAGTC 420
Db 361 GAAGCGGCTCTGCTGACTCTGGAGAGATGGTAGAGTTCTCAATACGCCAGTAGAGTC 420

QY 421 ATTGAAGAGAGCCTGTCAAAAGAGCTTCTGAACCTGGAGTGGTTACTCCAGGCAACTATAAC 480
Db 421 ATTGAAGAGAGCCTGTCAAAAGAGCTTCTGAACCTGGAGTGGTTACTCCAGGCAACTATAAC 480

QY 481 ACACCTGCACAAATCGTATCTCTGGAGAGTGGTTGAGTTGATCGAGCGGTTGAATTT 540
Db 481 ACACCTGCACAAATCGTATCTCTGGAGAGTGGTTGAGTTGATCGAGCGGTTGAATTT 540

QY 541 TTGCAAGAGAGAGGTCGCAAGAGCTTATTCCTCTTAAGGTGTCAGGTCCCTTTCACACC 600
Db 541 TTGCAAGAGAGAGGTCGCAAGAGCTTATTCCTCTTAAGGTGTCAGGTCCCTTTCACACC 600

QY 601 GCTCTCTTTGAGCCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAGTTTTC 660
Db 601 GCTCTCTTTGAGCCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAGTTTTC 660

QY 661 GATTTCATGTCCTCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 720
Db 661 GATTTCATGTCCTCTAGTCGGCAATACAGAGCTGCTGTGATGCAAAAAGAGACATT 720

QY 721 GCTCAGCTCTGACGCGTCAGGTCAAGGAACCGGTTCTGTTCTATGAAGATTATGGGTC 780
Db 721 GCTCAGCTCTGACGCGTCAGGTCAAGGAACCGGTTCTGTTCTATGAAGATTATGGGTC 780

QY 781 ATGCAAGAGAGAGGCAATAGCAACTTTATCGAGATTGACCGGGGAAAGTCTGTGAGGT 840
Db 781 ATGCAAGAGAGAGGCAATAGCAACTTTATCGAGATTGACCGGGGAAAGTCTGTGAGGT 840

QY 841 TTTGTTAAAAAATTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTA 900
Db 841 TTTGTTAAAAAATTGATCGAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTA 900

QY 901 GTAGCACTTTTGAAGAAATAG 921
Db 901 GTAGCACTTTTGAAGAAATAG 921

RESULT 2
US-09-815-242-6823
; Sequence 6823, Application US/09815242

Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Chlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA 011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6823
LENGTH: 942
TYPE: DNA
ORGANISM: Enterococcus faecalis
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(942)
US-09-815-242-6823

Query Match 37.4%; Score 344.4; DB 10; Length 942;
Best Local Similarity 61.3%; Pred. No. 4.5e-99;
Matches 555; Conservative 0; Mismatches 351; Indels 0; Gaps 0;

QY 7 AAAACAGCCTTTTATTTGCTGTCAGGTGCGGATCTAGGATCGGATTTTC 66
Db 16 AAAACAGCCTTTTATTTGCTGTCAGGTGCGGATCTAGGATCGGATTTTC 75

QY 67 TATGATCAGTATCCGATTCGAAAGAACGATTGATCGAGCGAGTCAGGTCTCGGTTAT 126
Db 76 TATCACCAGAGAGCGATTGTCGGGAACTTTCGATGAAGCAAGTCATCTTAGTTAT 135

QY 127 GATTTACGTTATCTCATCGATCGGAAAGACAACTCAATCAGACCGCTATACGCAA 186
Db 136 GAGATCGCAGAACTTTGTTTACTGAAATGAACGTTTAAATGAAACAGATATACGCAA 195

QY 187 CCAGCCATCTAGCGACTTCGTTGCTATCTACCGTTTATTCGAGAAAGGCGTATCAG 246
Db 196 CCGTCTATTTTAAACAGTCAGTTCGCAATTTTACCGCTTTTTCACAAAAGGACTAAACG 255

QY 247 CCTGATATGTTGCTGCTGTTGCTCTTGGAGATATCTCTGCTTGGCAAGCGCGCC 306
Db 256 CTTGATGTCGTCAGCGGTTTAACTTAGGGATATAGCGCTTGGTTCCAGCGGGCT 315

QY 307 TTGGATTTGAAGATCGGTTGCTTGGTAGCTAGCGTGGAGCCCTATATGGAAGAGCG 366
Db 316 TTGCGCTTTTTCAGAGCAGTGGCGCTTGGTCCAAAAGCGCGGTCAGTACATGACAGAGCA 375

QY 367 GCTCCTGCTGACTCTCGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAGGTCATTGAA 426
Db 376 GCACCAAGAGAACTGGCAAAATGGTTGCTGTCATGCAATGCTGAGCGTGAAGTAATTGAG 435

QY 427 GAAGCCTGTCAAAAAGCCTTCTGAACCTTGGAGTGGTTACTCCAGCCCAACTATAACACCT 486

436 AAACCTTCCAGAGCCAGTCTTTTCGGAATTTGGCTCCAGCAATTTATATACCA 495
487 GCACAAATCGTCAATGCTGGAGAGTGGTTCAGATTGATCGAGCGTGTGAACATTTTGC 546
496 CAACAAATCGTGAATGCTGGTGGTGGTCTGCTGTTGATCAAGCGATGACACTTCTCAA 555
547 GAACAGGTCCTCAACCGTTCATTCCTCTTAAGGTGTCAGTCCCTTTACACCGCTCTC 606
556 GAACCTGGTGAAGCGAATGATTCCTTAATGTAGTGGCCCTTTCCATACGGCGTG 615
607 CTGAGCGCTGTCAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAACTTTTTCAGATTT 666
616 TTACACAGCATCAAAAAATTTGGCTCAGGATTTAGCAAAATTTGAATTTCAACGATG 675
667 ACTTGTCCTAGTCGGCAATACGAAGCTGCTGTGATGCAAAAGAGACATTTGCTCAG 726
676 CAATTCCTGCTAATAGTAATAGCTGCGCAAAATTTATGCCCAAGAGCAATTTCAAGCG 735
727 CTCTTGACGGCTCAGGTCAAGGAACCGTTTCGTTTCTATGAAAGTATTTGGGGTCATGCAA 786
736 TTATTTGAAAGCAAGTCATGCTCGGTAGCTTTTGAAGACAGTATCGAAAGCATGAAG 795
787 GAACAGCGCATAGCAACTTTATGAGATTGGACCGGGAAAGTCTGTGACGTTTGT 846
796 GCTATGAACGTGGGAACGATGATTAAGTGGTCCAGGGGAAACATTAACGTGTTTGT 855
847 AAAAAATTCATCAACTGCTCACTAGCTCATGTGAGATCAAGCGAGTTTAGTAGCA 906
856 AAAAAATTCACAAAAATTAAGTGAATGACCGGTGAGATTTTGAACATTAACAGAA 915
907 CTTTTA 912
916 ACGTTA 921

RESULT 3

US-09-070-927A-577
; Sequence 577, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369

TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 577:
SEQUENCE CHARACTERISTICS:
LENGTH: 3656 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 577:
US-09-070-927A-577

Query Match 37.0%; Score 340.4; DB 10; Length 3656;
Best Local Similarity 61.7%; Pred. No. 2e-97;
Matches 556; Conservative 1; Mismatches 343; Indels 1; Gaps 1;

QY 7 AAACAGCGCTTTTATTTGCTGGTCCAGGTGCCAGTATCTAGGATGGGCGGATTTTC 66
Db 2753 AAACAGCGATTTTATTTAGTGGCAAGGAGCCAGTATCAAGGATGGGTGAAGATTA 2812
QY 67 TATGATCAGTATCGATTGTCAAGAAACGATTGATCGAGCGAGTCAGGTGCTCGGTTAT 126
Db 2813 TATCACCAGAGGCGATTTGTTCCGGAACTTTTCGATGAAGCAAGTCATATCTTAGGTTAT 2872
QY 127 GATTTACGTTTATCTCATCGATACGGAAGAGACAACTCAATCAGACCCGCTATACGCAA 186
Db 2873 GAGATGGCAGAACTTTGTTTACTGAAATGAACTTAAATGAAACAGATATACGCAA 2932
QY 187 CCAGCATTTAGGCACTTCGGTTGCTATCTACCGTTTATTCGAGAAAGAGGCTATCAG 246
Db 2933 CTTGATTTTAAACAGTCAGTGTGCGATTTTACCGTCTTTTTCGCAACAAAAAGGACTAACG 2992
QY 247 CTTGATATGTTGCTGGTTTCTCTTGGAGAACTCTGCTTGGTGGGAGCGGGGCC 306
Db 2993 CTTGATGCTGAGCGGTTTAAAGCTTAGGGAAATACAGTGTCTTTGGTCCAGCGGGGCT 3052
QY 307 TTGGATTTTGAAGATGCGGTTGCTTGGTGTAGCTAAGCGTGGAGCCTATATGGAAGAAGCG 366
Db 3053 TTGGCTTTTTCAGAAAGCAGTGGCTTGGTCCAAAGCGCGGTAGTACATGACAGAAAGCA 3112
QY 367 GCTCTGCTGACTCTGGCAGATGCTAGCAGTCTTCAATCGCCAGCAGTAGAGTCAATGAA 426
Db 3113 GCACCAAGGAACTGGCAAAATGGTTGCTGTTATGAATGTGAGCGTGAAGTAATTCAG 3172
QY 427 GAAGCCTGTCAAAAGGCTTCTGAACTTGGAGTGGTTACTCCAGCCAACTATAACACACCT 486
Db 3173 AAAGCTTGCCAAAGAGCAGTGGCTTTCGGAATTTGCTGGCTCCAGCAATTTATAATACCA 3232
QY 487 GCACAAATCGTCAATGCTGGAGAAAGTGGTTCAGTTGATCGAGCGGTGAACTTTGCAA 546
Db 3233 CAACAAATCGTGAATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3292
QY 547 GAAGCAGGTGCCAAACGCTTCTTAAAGTGTGAGTCCCTTTTACACCGCTCTC 606
Db 3293 GAAGCTGGTGGAGAGCAATGATTCGGTAAATGTGAGTGGCCCTTTCCATACCGCGCTG 3352
QY 607 CTTGAGCCTGTAGCCAGAAACTAGCTGAAACTCTAGCTCAGGTAAAGTGTTCAGATTTT 666
Db 3353 TTACAAACAGCATCAAAAAATTTGGCTCAGGATTTAGCAAAATTTGAACACTTTCAACAGATG 3412
QY 667 ACTTGTCCCTAGTGGCAATACAGAGCTGCTGTGATGCAAAAGAGGACATTCCTCAG 726
Db 3413 CAATTTCTGTCATTAGTATACGACTGCGCAATTTATGCCCAAGAGGCAATTTCAAGCG 3472
QY 727 CTTTGAGCGCTCAGGTCAAGGAAACCGGTTCTGTTTCTTATGAAAGTATTTGGGTCATGCAA 786
Db 3473 TTAATGGAAGCAAGTCATGCTGCGGTACGTTTTCAGACAGTATCGAAACGATGAAG 3532
QY 787 GAACAGGCATACCAACTTTATCG-AGATTGGACCGGGAAAGTCTGTGACGTTTTGT 845
Db 3533 GCTATGAACGTAGGAACGATGATTGAAGTNGTCCAGGGRAAACATTAACYGGTTTTGT 3592
QY 846 TAAAAAATTTGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAAGCGAGTTTAGTAGC 905

Db 3593 TAAAAAATGACCAAAATTAATGAATGACCGTGTGGAGATGTTGCCACATTAACAGA 3652
QY 906 A 906
Db 3653 A 3653

RESULT 4

US-09-974-300-1143
; Sequence 1143, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1143
; LENGTH: 950
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1143

Query Match 22.1%; Score 204; DB 10; Length 950;
Best Local Similarity 54.6%; Pred. No. 2.5e-54;
Matches 474; Conservative 0; Mismatches 385; Indels 9; Gaps 3;
QY 1 ATGACTAAACAGCCCTTTTATTTGCTGCTCAAGTGCCTAGTATCTAGGATGGACGG 60
Db 1 ATGGCAAGATTTGCTTTTATTTCCGGGCAAGGTTGCGACATATCGGATGGACAC 60
QY 61 GATTTCTATGATCAGTATCGGATTCGAAGAAACGATTCGAGCGAGTCAGTGCTC 120
Db 61 GAATTTGATGAAGAAACCGATGCGAAGAGATTTTGAAGACGGATCAACGTT 120
QY 121 GGTATGATTTAGTTATCTCA---TCGATACGGAAGACAACTCAATCAGACCCGC 177
Db 121 GATTCATGATCAGTATCGGATTCGAAGAAACGATTCGAGCGAGTCAGTGCTC 120
QY 238 GGCTATCAGCCCTGATGTTGCTGCTTGTCTCTTGGAGATATCTGCTTGGTGGCA 297
Db 241 GCATTAAGCGCATATTCGCGCAGGTACAGCCCTCGGCAATACAGCGCATTTGCTGCT 300
QY 298 ACGGGCCCTTGGATTTGAAGATCGGTTGCTTGGTATGTAACGTTGAGCGCTATATG 357
Db 301 GCGGCGCCCTTCTGTTTAAAGATCGGTTTATGCGCTCAGAAAGCGGGCGAATTCATG 360
QY 358 GAAGAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Db 361 AATGAAGCGCTCGCGGGGGAAGCGCGGATGCGGCGCATTCGCGCATGCGACGCGAG 420
QY 418 GTCAITGAAGAGCGCTGTCAAAAAGCTTCTGAACCTTGA---GTGGTTACTCCAGCCAA 474
Db 421 GCGCTGAAGAAAGTACGACGACAAATTTCCGAAGAAAGGAAACCTTGTTCAGCTCGCCAA 480
QY 475 TATACACACCTTGCACAAATCTGATTCGTTGAGAGAGTGTTCGATTTGATCGAGCGGTT 534
Db 481 TTGAATGCTCCCTGGGCAAAATCTGATCTCGGGAACAGCTAAAGCGGTGAGCTCGCTTCA 540
QY 535 GAACCTTTGCAAGAGCAGGTGCCAAACGCTTTGATTCCTCTTAAGGTGTCAAGTCCCTTT 594

Db 541 GAGCTTTCGAAAGAAAGGGCGCAAAACGCGGATTCCTCTCGAAGTCAGCGGCGCTTC 600
QY 595 CACACGCTCTCTTCTGAGCTGCTAGCCAGAAACTAGCTGAAACTCTAGCTCAGGTAAGT 654
Db 601 CATTTGAGCTGATGAAGCGGCGAGCTGATAGCTTCGTGAGTCTTGTATCGTGCAG 660
QY 655 TTTTCAGATTTTACTTGTCTCCCTAGTCGGCAA---TACAGAGCTCTCTGTGATGCAAAA 711
Db 661 ATCAAGCAGCAGCCATTCGGTCTCTCCAACTGTAACGCGGACTTTTAAACGGATAA 720
QY 712 GAGGACATTCCTCAGCTCTTTCAGCGCTCAGGTCAAGAAACCGTTCGTTTCTATGAAAGT 771
Db 721 GACGACATTAAGATAAATGATTTGACAGCTGATTCCTCCCTGTAGCTTTGAGAAACA 780
QY 772 ATTGGGTCATGCAAGACGAGCATATGAACTTTTATCGAGATTGACCGGGAAGTTC 831
Db 781 ATCAGCGCGCTGATTGACGAAGCGTCAGCACCTTCATTGAAATCGTCCCGAAAGTT 840
QY 832 TTGTCAGGTTTGTGTTAAAAAATTGATC 859
Db 841 TTGTCAGGCTTGTGAAGAAAGTGAACC 868

RESULT 5

US-09-974-300-1086
; Sequence 1086, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1086
; LENGTH: 960
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1086

Query Match 22.1%; Score 204; DB 10; Length 960;
Best Local Similarity 54.8%; Pred. No. 2.6e-54;
Matches 474; Conservative 0; Mismatches 385; Indels 9; Gaps 3;
QY 1 ATGACTAAACAGCCCTTTTATTTGCTGCTCAAGTGCCTAGTATCTAGGATGGACGG 60
Db 1 ATGGCAAGATTTGCTTTTATTTCCGGGCAAGGTTGCGACATATCGGATGGACAC 60
QY 61 GATTTCTATGATCAGTATCGGATTCGAAGAAACGATTCGAGCGAGTCAGTGCTC 120
Db 61 GAATTTGATGAAGAAACCGAATGCGAAGAGATTTTGAAGAAGCGGATCAAAAGCTT 120
QY 121 GGTATGATTTAGTTATCTCA---TCGATACGGAAGACAACTCAATCAGACCCGC 177
Db 121 GAACAAACTGACACCCCTCATGTTGAAGGGATGCAAGGAAGTACGCTTACATAC 180
QY 178 TATACGCAACCGACCTTCTAGCGACTTCGTTGCTGTTGCTCTCTTGGAGATATCTCTCCCTTGGTGGCA 237
Db 181 ACGCGCAGCAAGCCCTTTTAAACGCGGAGCATCGCAGCGCTTGAAAAAAGTGAAGAAAT 240
QY 238 GGCTATCAGCTGATATGTTGCTGCTGTTGCTCTCTTGGAGATATCTCTCCCTTGGTGGCA 297
Db 241 GGCATTAAGCGCATATTCGCGCAGGTACAGCCCTCGGCAATACAGCGCATTTGCTGCT 300
QY 298 ACGGGCCCTTGGATTTGAAGATCGGTTGCTTGGTATGTAACGTTGAGCGCTATATG 357
Db 301 GCGGCGCCCTTCTGTTTAAAGATCGGTTTATGCGCTCAGAAAGCGGGCGAATTCATG 360
QY 358 GAAGAAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Db 361 AATGAAGCGCTCGCGGGGGAAGCGCGGATGCGGCGCATTCGCGCATGCGACGCGAG 420
QY 418 GTCAITGAAGAGCGCTGTCAAAAAGCTTCTGAACCTTGA---GTGGTTACTCCAGCCAA 474
Db 421 GCGCTGAAGAAAGTACGACGACAAATTTCCGAAGAAAGGAAACCTTGTTCAGCTCGCCAA 480
QY 475 TATACACACCTTGCACAAATCTGATTCGTTGAGAGAGTGTTCGATTTGATCGAGCGGTT 534
Db 481 TTGAATGCTCCCTGGGCAAAATCTGATCTCGGGAACAGCTAAAGCGGTGAGCTCGCTTCA 540
QY 535 GAACCTTTGCAAGAGCAGGTGCCAAACGCTTTGATTCCTCTTAAGGTGTCAAGTCCCTTT 594

```

Db 301 GCCGGCCCTTGTCTTAAAGATCGGTTTATGCCGTGAGAAGCCGGCGAATTCATG 360
Qy 358 GAAGAAGCGGCTCTGTGACTCTGGCAAGATGTAGCAGTTCTCAATACGCCAGTAGAG 417
Db 361 AATGAAGCCGTGCGGGGAGAGGCGGATGCGGCCATTCTCGCATGGACAGCCAG 420
Qy 418 GTCAATCAAGAAGCCTCTCAAAAGCTTCTGAACCTTGA---GTGTTTACTCCAGCAAC 474
Db 421 GCGCTGAAGAAGTAGCGGCAAAATTTCCGAAGAAGAAACCTTGTTCAGCTCGCCAAT 480
Qy 475 TATAACACACCTGCACAAATCGTCATTCTCGGAGAAGTGTTCGAGTTGATCGAGCGGTT 534
Db 481 TTGAACCTGCGCTGGCAAAATCGTCATCTCGGAACAGCTTAAGCGGTGAGCTCGCTTCA 540
Qy 535 GAACCTTTTGAAGAAGCAGGTGCAACAGCTTGAATCTCTTAAGGTGTGAGTCCCTTT 594
Db 541 GAGCTTCGGAAGAAAGGCGCAAAACGCGCGAATTCCTCTCGAAGTCAAGCGGCGGTT 600
Qy 595 CACACCCCTCTCTCTGAGCCTGCTAGCCAGAACTAGCTGAAACTCTAGCTCAGGTAAGT 654
Db 601 CATCTGAGCTGATGAAGCGGAGCTGATGAAGCTTCGTGAAGTTCTTGAATGCTGACG 660
Qy 655 TTTTCAGATTTTACTTGTCCCTAGTGGCAA---TACAGAAGCTGCTGTGATGCAAAA 711
Db 661 ATCAACGACGAGCATTCCGCTCGTCTCCAAGCTAAGCGGCACTTTGTACGGATAAA 720
Qy 712 GAGGACATTCCTCAGCTCTTGAAGCGCTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 771
Db 721 GACGACATTAAGAATAAAGTATGAAGCTGATTTCCCTCTGACGCTTTGAAGAAACA 780
Qy 772 ATTGGGTCATGCAAGAAGCAGGATGAAGCAACTTTATCGAGATTGAGCCGGGAAAGTC 831
Db 781 ATCAGCCGCTGATTGACGAAGCGCTCAGCAGCTTCATTGAAATCGTCCCGGAAAGTT 840
Qy 832 TTCTCAGCTTTTCTTAAAAAATGATC 859
Db 841 TTGTGAGGCTTGTGAAGAAGTGAACC 868

```

RESULT 6

```

US-09-815-242-6878
; Sequence 6878, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITEA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6878
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(939)
US-09-815-242-6878

```

```

Query Match      21.3%; Score 196.4; DB 10; Length 939;
Best Local Similarity 54.1%; Pred. No. 6.7e-52;
Matches 470; Conservative 0; Mismatches 386; Indels 12; Gaps 3;

Qy 1 ATGACTAAACAGCCCTTTTATTTGCTGCTCAAGTGCCCGCAGTATCTAGGATGGAGCG 60
Db 1 ATGAAAAATTCGCAATGGTCTTCCAGGTCAAGGCTCCCAAACTGTGGTGTGCTTCT 60
Qy 61 GATTTCTATGATCAGTATCCGATTTGTCAAAGAAACGATTTGATCGAGCGAGTCAGTGCTC 120
Db 61 GATCTTGAACACTGATATCCAAATCGTTATTGAAACATTTAAACAAGCATCTGTATGCGTT 120
Qy 121 GGTATGATTTACGTTATCTCATCGATACGG---AAGAAGACAAACTCAATCAGACCCGC 177
Db 121 GGTATGATTTATGGTATCTTTGTTCAACAAGGTCCAGCTGAAGAACTTAATAAAACCTTG 180
Qy 178 TATAGCAACCCAGCATTCTAGCAGCTTCGTTGCTATCTACCGTTTATTCGCAAGAAAG 237
Db 181 CAACTCAGCCCGCACATTTTATGCTGCTTATGCTGCTATTTATATCGCTATGGAAGAAAA 240
Qy 238 GGCATC-----AGCCTGATATGTTGCTGTTTGTCTCTTTGGAGAACTACTCTGCCTTG 291
Db 241 TTTCTCAATTAATAAACAGAAAGTATGCGAGTCTAGCTTAGTGTAGTATTTCTGCGTTA 300
Qy 292 GTGCAAGCGCGCTTGGATTTTGAAGATCGGTTGCTTGGTGTAGCTAAGCGTGGAGCC 351
Db 301 GTTTGTGCTGCGGTGTTGATTTCCAAAGATGCGAATTAATTAGTGGAAATTCGCGGAAAA 360
Qy 352 TATATGAAGAAGCGGCTCTGCTGACTCTGGCAAGATGTTAGCAGTTCTCAATACGCCA 411
Db 361 TTAATGCAACAAGCTGTGCTCTGAAGCACTGCGCAATGTATGCAATCATTTGTTAGAT 420
Qy 412 GTAGAGTCATGAAGAAGCCTGTCAAAAGCTTCTGAACTTGGAGTGTGTACTCCAGCC 471
Db 421 AATGAAGCAATTTAATATGCTTTCAAAACAAGCAGGAAGCGAAGTCGTATCTGCGGTG 480
Qy 472 AACTATACACACCTGCACAAATCGTCATTCTGAGAGAGTGTTCAGATTGATCGAGCG 531
Db 481 AACTTTAACTCAGCGGTCAAGTAGTTATTGCGGTTGCGAAGCTGAGTTGAGCGTGG 540
Qy 532 GTTGAACCTTTTGAAGAAGCAGGTGCCAAAGCTTTGATTCCTCTTAAGGTGTTCAGGTCCC 591
Db 541 GCTGCATTTATGTAAGAAGCAGGCGGCAAGCTGCAATTCGCTTAGCTGTAGCGGTACCT 600
Qy 592 TTTCAACCGCTCTCTTGTAGCCTGTAGCCAGAACTAGCTGAACCTTAGCTCAGGT- 650
Db 601 TCTCAGTGTGCATTAATGAACCTGAGCGGAGCAATTAGCGGTAAACATTTGAAATATT 660
Qy 651 --AAGTTTTTTCAGATTTTACTTTGCTCCCTAGTCGGCAATACAGAAGCTGTGTGATCAA 708
Db 661 CAAATTAATACCAACAATATCGGTATTAAATAACCTTGTATGTAAGCTGAAACTGAA 720
Qy 709 AAGAGGACATTTGCTCAGCTTTGACCGGTGAGTCAAGGAAACCCGTTTGGTTCTATGAA 768
Db 721 GGCACCGAAATTCGTACCGCACATTTGCTGCTGCTATATAGTCAGTTCGTTGAGCTGAA 780
Qy 769 AGTATTGGGTCATGCAAGAAGCAGGATAAGCAACTTTTATCGAGATTGACCGGGGAAA 828
Db 781 ACAGTTGAAGAAATGGCGCAGATGGCGTCTAGTGTCTGCTGAGTGGCGCCAGGTAAA 840
Qy 829 GTCCTGTCAGGTTTGTAAAAAATTTG 856
Db 841 GTATTAATGTTTAAACCAACGCAATTG 868

```

```
RESULT 7
US-10-329-960-1/c
; Sequence 1, Application US/10329960
; Publication No. US20030099277A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Theorof, and Uses Thereof
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,960
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
```

```

NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152530)..(152530)

```

```

Query Match 21.3%; Score 196.4; DB 9; Length 1830121;
Best Local Similarity 54.1%; Pred. No. 7.5e-50;
Matches 470; Conservative 0; Mismatches 386; Indels 12; Gaps 3;

QY 1 ATGACTAAACAGCGCTTTTATTTGCTCAAGTGCCCGAGTATCTAGGGATGGGACGG 60
DB 172507 ATGAAAATTCGAATGGCTCCCGAGTCAAGGCTCCCAACTGTCGGTATGCTTGT 172448
QY 61 GATTCTATGATCATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 120
DB 172447 GATCTGCACTGAATATCCCAATCGTTATTCGAACATTTAAACAAGCATCTGATGCGCT 172388
QY 121 GGTATGATTTACCTTATCATGATACGG---AAGAAGACAACACTCATGACCCGC 177

```

```

Db 172387 GGTATGATTTATGTTATCTTGTTCACAGGTCAGCTGAAGAACTTAATAAACTTGG 172328
QY 178 TATACGCAACAGCCATTCTAGCGATTTCGGTTGCTATCTACCGTTTATTCGAAGAAAG 237
Db 172327 CAAACTCAGCCGCGCACTTTTAGCTGCTTTCAGTCTGCTATTTATCGCTATGGAAGAAAA 172268
QY 238 GGCTATC-----AGCCTGATATGGTTGCTGTTTGTCTCTTGGAGAACTCTGCTTGG 291
Db 172267 TTTCTCAATTAACCAAGAGTATGGCAGGTCTAGCTTAGTGTAGTATTTCTGGCTTA 172208
QY 292 GTGGCAAGCGCGCTTTGGATTTTGAAGATCGGGTTGCCCTTGTAGTAAAGCGTGGAGCC 351
Db 172207 GTTTGTGCTGGCGTGTGGATTTCGAAGATCGGATTAATTAATAGTGAAATTCGCGGAAAA 172148
QY 352 TATATGGAAGAGCGGCTCTCTGCTGCTGCTGCAAGATGGTAGCAGTTCTCAATACGCCA 411
Db 172147 TTAATGCAACAAGCTGTGCTCTGAAGGCACCTGCGCAATGATGCAATCAATGGTTTAGAT 172088
QY 412 GTAGAGGTCAATTGAAGAGCCTGTCAAAAAGCTTCTGAACTTTGGAGTGGTTACTTCAGCC 471
Db 172087 AATGAAGCAATTTAATGCTTTGCAACAAGCAGAGGAGCGCAAGTCTGATCTCGGTG 172028
QY 472 AACTATACACACCTGCAAAATCGTCAATGCTGAGAGAGTGGTTGCAAGTTCGATTCAGCG 531
Db 172027 AACTTTAACTACCGGTCAGTAGTATTGCGGGTGGAAAAAGCTGCAAGTTGAGCGTGG 171968
QY 532 GTTGAACCTTTTGAAGAGCAGTGTCCAAACGCTTGTATTCCTTAAAGTGTGAGTCCC 591
Db 171967 GCTGCAATTTATGAAGAGCAGCGGCGGAAACGTCGATTCGCTGAGCGTACCT 171908
QY 592 TTTACACCGCTCTCTTGGAGCTGCTAGCCAGAACTAGCTGAAACTTACTCAGGT- 650
Db 171907 TCTCACTGTGCATTAATGAAACCTGACGCGGCAATAGCGGTAAACCTGAGAATAT 171848
QY 651 --AAGTTTTCAGATTTTACTTGTCCCTAGTCGCGCAATACAGAGCTGCTGTGATGAA 708
Db 171847 CAATTAATACCAACCAATATCGGTATTAAATACGTTGATGGAAGCTGAAACTGAA 171788
QY 709 AAAGAGCAATTTGCTCAGCTCTTACCGCTCAGTCAAGGAAACCGCTTCGTTTCTATGAA 768
Db 171787 GGCACCGAAATTCGTACCGCACTTGTGCGTCAAGTATATAGTCCAGTTCGTTGGACTGAA 171728
QY 769 AGTATGGGTCATGCAAGAGCAGCATAGCACTTATCGAGATTGGACCGGGGAA 828
Db 171727 ACAGTTGAAAAATGGCGCAAGATGCGCTTCTAGTGTGTAAGTGGGCCAGGTAAA 171668
QY 829 GTCTTGTGAGGTTTGTGTTAAAAAAATTG 856
Db 171667 GTATTAAATGTTTAAACCAACGCAATTG 171640

```

```

RESULT 8
US-09-815-242-6028
; Sequence 6028, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23

```

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6028
LENGTH: 930
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(930)
US-09:815-242-6028

Query Match
Best Local Similarity 20.2%; Score 185.6; DB 10; Length 930;
Matches 464; Conservative 0; Mismatches 399; Indels 9; Gaps 3;

QY 1 ATGACTAAACAGCCCTTTTATTTGCTGTCAGGTCGCCAGTATCTAGGGATGGGACGG 60
DB 1 ATGACGCAATTGCAATTTGTTTCCCTGGACAGGGTTTCAAAACCGTTGGAATGCTGGCT 60
QY 61 GATTCTATGATCAGTATCGATGTCGCAAGAACGATGATCGACGAGTCAAGTGCTC 120
DB 61 GATATGGCGGAGCTATCCAAATGTGCAAGAACGATTTGCTGGAAGCTTCTCGCGGCTG 120
QY 121 GGTATGATTTAGCTATCTCATCGATACGGA---AGAAGACAACTCAATCAGACCCGC 177
DB 121 GGCTAGACCTGCGGCGTACCCAGCAGGSCAGCTGAGCACTGATTAACCTGG 180
QY 178 TATACGCAACAGCCATTTAGGACCTGCTGGTTGCTATCTACCGTTTATGCAAGAAAG 237
DB 181 CAAACTCAGCTCGCTGTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 238 GGCTATCA---GCTGATATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
DB 241 GCGGTAAGCAGCCGCAATGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 295 GCAAGCGGCGCTTGAATTTGAAGATGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
DB 301 TGGCTGGTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 355 ATGGAAGAGCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
DB 361 ATGCAAGAGCCTAATCGGAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 415 GAGGTCAATGAAGAGCCTTCAAAAAGCTTCTGAACTTGGAGTGGTTACTCCAGCAAC 474
DB 421 GCGTCTATTGCGAAGCGCTGGAAGAGTGCAGAGGTCAGTCTGCTTCTCCGGTAAAC 480
QY 475 TATAACACACTGCACAAATCGTCATTTGCTGGAGAGTGGTTCAGTTCATCGAGCGTT 534
DB 481 TTTAACTCTCCGGAGAGGGTGTATTGCGGCTATTAAGAGCGGTTGAGCGTCTGGC 540
QY 535 GAACTTTTGAAGAGCAGTGCACAAACGCTTGAATTCCTTCTTAAAGTGTGAGTCCCTTT 594
DB 541 GCTGCTGTAAAGCGGCGGCGCAAAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 595 CACACGCTCTCTGAGCTGTAGCAGAACTAGCTGAACTCTAGCTAGCTAGCTAGCTAGT 654
DB 601 CACTGTGCGTGTGAACACAGCAGCGCAAACTGGCAGTAGAATTAGCGAAATCACC 660
QY 655 TTTTTCAGATTTTACTTGTCCCTAGTCGGCAATACAGAAAG---CTGCTGTGATGCAAAA 711
DB 661 TTTAACGCCACACAGTTCTCTGTTGTGATACGTTGATGTGATGCAATGCAAAACCATGGT 720

QY 712 GAGGACATTTGCTCAGCTCTTGAACGCTCAGGTCAAGGAACCCGTTCTTCTATGAAAGT 771
DB 721 GATCCATCCGTGACGCACTGGTACGTCACTTGTATTAACCCGTTCACTGACGAGTCT 780
QY 772 ATTGGGTCATGCAAGAGCAGGATAGCAACTTTATCGAGATTGGACCGGGAAGTC 831
DB 781 GTTGAGTACATGGCAGCGCAGGCTAGACATCTCTATGAGTCGCGCCGCGCAAGTG 840
QY 832 TTGTCAGGTTTGTAAAAAATTTGATCAAAAC 863
DB 841 CTTACTGCGCTGACGAAACGCAATTTGCGACAC 872

RESULT 9
US-09-070-927A-458
Sequence 458, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Patrick J. Dillon
Steven Barash
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB369
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 458:
SEQUENCE CHARACTERISTICS:
LENGTH: 6021 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 458:
US-09-070-927A-458

Query Match
Best Local Similarity 17.2%; Score 158; DB 10; Length 6021;
Matches 266; Conservative 4; Mismatches 184; Indels 0; Gaps 0;

QY 459 GGTACTCCAGCAACTATAACACACTGCAAAATCGTCATTGCTGGAGAGTGGTTC 518
DB 172 GGTGCTCCAGCAAAATATATWATWCMCMCAAAATCGTGATTGGTGGTGGTTCGTC 231
QY 519 AGTGTATCGAGCGGTTGAACCTTTTGAAGAGCAGCTGCCAACGCTTGATTCCTTAA 578
DB 232 TGTTGATCAAGCGGATGACACTTCTCAAAGAAAGTGGTGTGAGCGAATGATTCGTTAA 291

```
QY 579 GGTGTAGTCCCTTTTACACCGCTCTCTTGTAGCCCTGCTAGCCAGAAACTAGCTGAAC 638
DB 292 TGTAGTGGCCCTTTTCCATACGGGCTGTTACAAACAGCAGCATCAAAAAAATTGGCTCAGGA 351
QY 639 TCTAGCTCAGTAAGTTTTTTCAGATTTTACTTGTCCCTAGTGGCAATACAGAAGCTGC 698
DB 352 TTTAGCAAAATTGAACTTTCAACAGATGCAAAATTCCTGTATAGTAAATAGACTGCCGA 411
QY 699 TGTGATGCAAAAGAGGACATTTGTCTAGCTCTTGACCGCTCAGGTCAAGGAAACCGGTG 758
DB 412 AATTATGCCCCAAGAGGCAATTAAGCGTTTATTGAAAAAGCAAGTCATGCTCGGCTACG 471
QY 759 TTTCTATGAAAGTATTTGGGTCATGCAAGAGCAGGAGGATAGCAACTTTATCGAGATTGG 818
DB 472 TTTTGAAGACAGTATGAAACGATGAAGGCTATGAAGCTAGGAAACGATGATTAAGATTGG 531
QY 819 ACCGGGAAAGCTTTGTGTCAGCTTTTGTATAAAAAATTGATCAAACTGCTCACTTAGCTCA 878
DB 532 TCCAGGAAAAACATTAACCTGTTTGTATAAAAAATTGACAAAAAATTTGAATGACCG 591
QY 879 TGTGAAGATCAAGCGAGTTTAGTACACTTTTA 912
DB 592 TGTGAAGATGTTGCAACATTAACAGAAACGTTA 625

RESULT 10
US-09-974-300-5552
; Sequence 5552, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/974,300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5552
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-5552

Query Match 15.2%; Score 139.6; DB 10; Length 592;
Best Local Similarity 54.8%; Pred. No. 6.3e-34;
Matches 321; Conservative 0; Mismatches 259; Indels 6; Gaps 2;

QY 2 TGACTAAACAGCCTTTTATTTCTGTCGAAGTGCCCAAGTATCTAGGGATGGGACGG 61
DB 2 TGGGAAAGTTGCCCTTTTATTTCTCGACAGGCTCAAAAAGTTGGATGGGAGCAA 61
QY 62 ATTTCTATGATCATCCGATTCGAAAGAAAGATGATCGAGCGAGTCAGTGCTCG 121
DB 62 CGCTTTTGTGTAAGATTTGATGCAAGGAAGTGTGAGCGGCTGACCAAGCGCTTG 121
QY 122 GTTATGATTTACGTTACTCTATCG ---ATACGGNAGAGCAAACTCAATCAGACCGCT 178
DB 122 GTTTTCTCTTTACAGATCATCGAATCGAACATGGACCAGAGACAGCTGAACAAACCGCTT 181
QY 179 ATAGCAACACAGCACTTCTAGCGACTTCGGTTGCTATCTACCTGTTTATTCGAAGAAAG 238
DB 182 ATGCCCAACCTGCTTTTAGTGAGGATGAGCACTGCGCTCTCAATGTTTCCGGACGCTG 241
QY 239 GCTATACGCTGATATGGTTGCTGGTTTGTCTCTTGAGAAATACTCTGCTGTTGGTGA 298
DB 242 ACATTCAGCAGATTTTGTGTCAGGCCATACGCTAGGGGAGTATTCGGCTTTAGTGGCAT 301
```

```
QY 299 GCGGCGCCTTGGATTTTGAAGATGCGGTTGCTTGGTAGCTAAGCGTGGAGCCTATATGG 358
DB 302 GTAAGTCCCTTACATTTTGAAGATGCGGTACACTTGTTCATCAACGGGGACACTCATGG 361
QY 359 AAGAAGCGGCTCTCTGCTGACTCTGCAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGG 418
DB 362 AGAAGCTGTGCGCAAGACAGCAAGGGGCGATGGACGGTACTCGGCTTAATAAAGAGG 421
QY 419 TCATTGAAGAA---GCCTGTCAAAAAGCTTCTGAACITTTGGAGTGGTTACTCCAGCCAAC 475
DB 422 AGCTTGAGGAAGTGGCTTCCGAAATTTGACGGGACGGCGAAGTTGCTGAACTTGCCAATT 481
QY 476 ATAAACACACCTGCACAAATCGTTCATTGCTGGAGAGTGGTTGCAGTTGATCGAGCGGTTG 535
DB 482 TGAATTTGCCAGGGCAAAATTTGCTGTGTCAGGACATGCGAAAGGAATTTGAAACAGCAGG 541
QY 536 AACTTTTCAAGAACAGGTCGCAAAACGCTTGATTCCTCTCTTAAGGT 581
DB 542 TATTAGCAAAACAAAAGGGCCCAAGCGTGTCTTGCCACTTGCCGT 587

RESULT 11
US-09-815-242-8032
; Sequence 8032, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: Prokaryotes
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8032
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)....(936)
US-09-815-242-8032

Query Match 13.9%; Score 128; DB 10; Length 936;
Best Local Similarity 49.2%; Pred. No. 4.2e-30;
Matches 456; Conservative 0; Mismatches 455; Indels 15; Gaps 4;

QY 1 ATGACTAAACAGCCTTTTATTTCTGTCGAAGTGCCCAAGTATCTAGGATGGGACGG 60
DB 10 ATGACTAAACAGCAATTTATTTTCCGGACAGGTGCCCAAAAGTTGGTATGGCAA 69
```

Qy	61	GATTTCTATGATCAGTATCCGATTTGTCAAAGAAAGATGTATCGAGCGATGTCAAGTGGCTC	12
Db	70	GATTTTATAAACAACAACGATCAAGCAACTGAAATTTTAACTTCAGCAGCAAAACAGTTA	129
Qy	121	GGTTATGATTTAGCTTATCTCATCGATACGGAGAGA--CACAACCTCAATCAGACCCGC	177
Db	130	GACTTTGATATTTTAGAGACAATGTTTACTGACGAGACGGCAAAATTTGGGTGAACCTGAA	189
Qy	178	TATACGCAACACGCCATTTCTAGCGACTTCGGTTGCTATCTACCGGTTTATTTCGAAGAAAAG	237
Db	190	AACACGCAACACAGCTTTATTAGCGCATAGCTCGGGGTTA-----TTAGCGGGCTAAAA	243
Qy	238	GGCTATCAGCGCTGATAGTGGTTGCTGTTGCTCTTGGAGAACTACTGCTGCTTTGGTGGCA	297
Db	244	AAATTGAATCCCGATTTTACTATGGGGCATAGTTTAGTGAATATTCAAGTTTAGTTGTGA	303
Qy	298	AGCGGGCCCTTGGATTTTGAAGATCGGGTTGCCCTTGGTAGCTAGACGTGGAGCGCTATATG	357
Db	304	GCTGACGTATTATCATTTTGAAGATCGAGTTAAATTTGTAGAAAACGTGTCATTAATG	363
Qy	358	GAAGAACGGCTCTGCTGACTCTGGCAAGATGTTAGCAGTTCTCAATACGCCAGTAGAG	417
Db	364	GCGCAAGCATTTTCTACTGTTGTAGGAAGACATGCGTCAGTATTTGGCTTTAGATTTTGAT	423
Qy	418	GTCAATTGAAGAAAGCCGTGTC---AAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGCGCAAC	474
Db	424	AAAGTCGATGAAATTTTGTAAAGTCATTATCACTGATGACAAATAATTGAACCCAGCAAC	483
Qy	475	TATAACACACTGTCACAAATCGTCAATCTGCTGGAGAGTGGTTGCGATGTATCGAGCGGTT	534
Db	484	ATTAATTTGCCAGAGTCAAATTTGTTGTTTTCAGGTACACAAAGCTTTAAATTGATGAGCTAGTA	543
Qy	535	GAACTTTTTCGAAGACGAGTGCCTCAACGCTTGATTCCTCTTAAGGTGTCAGTCCCTTT	594
Db	544	GAATAAGGTAAATCATTAGTGCATAAGCGTGCATGCCCTTTAGCAGTATCTGGCCCATTC	603
Qy	595	CACACGCTCTCCTTTAGCGCTGCTAGCCAGAAACTAGCTGAAACTCTAGCTCAGTAAAGT	654
Db	604	CATTATCGCTAATGAAAGTGATTCAAGAAGATTTTCAAGTTATATTAATCAATTTGAA	663
Qy	655	TTTTAGATTTTACTGTGTCCCTAGTCGGCAATACAGAAGCT---GCTGTGATGCAAAA	711
Db	664	TGCGATGATGCTAAGTTTCCCTGTAGTTCAAAATGTAAATGCGCAAGGTGAACCTGACAA	723
Qy	712	GAGGACATTTGCTCAGCTCTTGACGGGTGAGGTCAAGGAAACCGGTTCTGTTTCTATGAAGT	771
Db	724	GAAGTAATTAATCTAAATATGTCAAAACAATATATTCACGAGTACAAATTCATTACTCA	783
Qy	772	ATTGGGGTCATGCAAGACGAGCATAGCAACTTTTATCGAGATGTGAACCGGGGAAAGTC	831
Db	784	ACAGAAATGGCTAATAGACCAAGGTGTTGATCATTTTATTGAAATTTGCTCTCGGAAAAGTT	843
Qy	832	TTGTCAGGTTTTGTTAAAAAAATTCATCAAACTGCTCACTTAGTCTCATGTGGAAGATCAA	891
Db	844	TTGCTCGGCTTAATTAATAAATAAATAATAGAGATTTAGTTAATCAATTCAAACTTTA	903
Qy	892	GGAGTTTAGTACACTTTTAGAAA	917
Db	904	GAAGATGTCGAAGGATGGAATGAAA	929

RESULT 12

```

RESULTS 12
US-09-815-242-4405
; Patent 4405, Application US/09815242
; Sequence No. US20020061566A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

```


Db 595 CATTTCATCGCTAATGAAGAGTCTGATGAAGAGATTTTCAAGTTATATTAATCAATTTGAA 654
QY 655 TTTTCAGATTTTACTTGTCCCTAGTGGCAATACAGAGCT---GCTGTGATGCCAAAAA 711
Db 655 TGGGATGATGAAGTTTCCCTGTAGTTCAAAATGTAATGCGCAAGGTGAATCTGACAAA 714
QY 712 GAGGACATTCGTGAGCTCTTGACGGGTGAGGTCAAGGAAACCGTTTCGTTTCTATGAAAGT 771
Db 715 GAAGTAATTAATCTAATATGTCACAAATTAATATTACCAAGTACAAATTCATTAATCTCA 774
QY 772 ATTTGGGTGATGCAAGAGAGGAGGATAGCAATTTTATCGAGATGACCGGGGAAAGTC 831
Db 775 ACAGAAATGGCTAATAGACCAAGGTGTTGATCATTTTATGAAATGCTCTGGAAGATT 834
QY 832 TTGTCAGGTTTTTGTAAAAAATTTGATCAAACTGCTCACTTAC 875
Db 835 TTGCTGGCTTAATTAATAAATAATAGAGATTTAAGTTTAAAC 878

RESULT 13

US-09-815-242-7796
; Sequence 7796, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/205,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7796
; LENGTH: 939
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(939)
US-09-815-242-7796

Query Match 13.18; Score 120.8; DB 10; Length 939;
Best Local Similarity 49.08; Pred. No. 8.3e-28;
Matches 410; Conservative 0; Mismatches 417; Indels 9; Gaps 3;
QY 13 GCCTTTTATTTGCTGCTCAAGGTGCGGATGCTAGGATGCGGATTTCTATGAT 72
Db 16 GCAATTCGCTCCCTGCCAGGTTGCAATCCCTCGCATGCTGCCAGTGGCGCC 75
QY 73 CAGTATCCGATTTCAAGAAACGATTCGAGCGAGTCAGGTGCTCGTTATGATTTA 132

Db 76 CAGCAGCGCTGGTGGCGGATACCTTCGCGAGGCGCTCGAGGCGCTCGGTTACGACCTT 135
QY 133 CGTTATCTCATCGATACGG---AAGAAGACAAATCAATACAGACCGCTATACGCAACA 189
Db 136 TGGGCGCTGGTCCAGATGGTCTGAAGAGCGCTGAACACGACGACAAAGACCCAGCG 195
QY 190 GCATTCTAGCAGCTTCGGTTGCTATCTACCGTTTATTGCAAGAAAGGGCTAT--CAG 246
Db 196 GCATCTCTTACGGTTTCGATCGGCTCTGGCGCTCTGGCTGGCGAGGGCGGTGGCGC 255
QY 247 CTGTATATGGTTGCTGGTTTGTCTCTTGAGAGAAATCTTGTCTTGGTGGCAAGCGCGC 306
Db 256 CGGCGCTTCTGCGCGGACAGCTTGGCGCAATATTCGCGCTGCTCGCGCCGAAAGC 315
QY 307 TTGGATTTTGAAGATGCGGTTGCTTGGTAGCTAAGCGTGAGCCCTATATGAGAAAGCG 366
Db 316 CTGGGCTTCCCGATGCGGTCAAGCTGCTGAGCGTAGGGCCAACTGATGACGAGCGC 375
QY 367 GCTCTGCTGACTCTGGCAAGATGGTAGCTTCTCAATACGCCAGTAGAGGTCTATTGAA 426
Db 376 GTTCCGGCGGCGAGGCGGATGCGCGCATCTTGGCTTGAAGACGCGATGATTG 435
QY 427 GAAGCTGTCAAAAAGCTTCTGAATCTGAGTGGTGTACTCCAGCCCACTATTAACACCT 486
Db 436 GCGGCTGTGCGGAGGCGGCGCCAGGCGAGGTGGTCAAGCGGTCAATTTCAACGCGCG 495
QY 487 GCACAAATCGTCATTGCTGGAGAGTGGTTGCAGTTGATCGAGCGGTGAACCTTTTGCAA 546
Db 496 GGGCAGGTAGTATCGCGGTGCGCGCTGCGCTGAGCGTGCCATCGAGGCATGCAAG 555
QY 547 GAAGCAGGTGCGCAAGCGTTGATCTCTTAAAGTGTGAGTCCCTTTTACACCGCTCTC 606
Db 556 GCACGCGCGCAAGCGCGGTGGCTTGCAGTCAAGCTGCGCTGCGCTGCGCTTGGCACTG 615
QY 607 CTTGAGCTGTGAGCCAGAACTAGCTGAACTCTAGCTCAGGTAAGTTTTTCAGATTT 666
Db 616 ATGCTCGCGCGCGAGAGCTTCCGCGCTCGTCAAGAGCTGCGTCAAGAGGAGGAGCATT 675
QY 667 ACTTGTCCCTAGTCCGCAATACAGAGCTGCTGTGATGCAAAAAGAGGACATTCCTCAG 726
Db 676 AAGATTTGCTGGTGCAGAACTGAGCGCGCTGCGCTGATCTCGATACGCTGCGC 735
QY 727 CTCT---TGACGCGTCAAGTCAAGCAACCGCTGCTTCTATGAAGTATTTGGGCTCATG 783
Db 736 CCGGACCTGTGGCAGCTGTACAGCCCGTTCGCTGGTGGAGAGCATCCAGCTGCTG 795
QY 784 CAAGAAGCAGGCATGAAGCAACTTTATCGAGATTGACCGGGGAAAGTCTTGTGAGG 839
Db 796 GCGGAAAGGGGCTCACCGAGCTGCTGAGTGGCGGCGCGCAAGGTCTCTGGCAGG 851

RESULT 14

US-09-781-986A-25
; Sequence 25, Application US/09781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A

```

; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6251 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-25

Query Match      12.9%; Score 118.4; DB 7; Length 6251;
Best Local Similarity 48.8%; Pred. No. 1.6e-26;
Matches 450; Conservative 0; Mismatches 461; Indels 15; Gaps 4;

QY 1 ATGACTAAACAGCCCTTTTATTTGCTGCTCAAGTGCCTCAGTATCTAGGGATGGGACGG 60
Db 3386 ATGAGTAAACAGCAATTTATTTCCGGGACAGAGTGCCTCAAAAAGTTGGTATGGCAA 3445

QY 61 GATTCTATGATCAGTATCGGATGTCAAGAAACGATTGATCGAGCGAGTCAGTGGTCT 120
Db 3446 GATTCTTTAAACAAATGATCAAGCAACTGAAATTTTAACTTCAGCAGCGAACACATTA 3505

QY 121 GGTATGATTAGCTTATCTCATCGATAGCGAAGAGAC---AAACTCAATCAGACCCGC 177
Db 3506 GACTTTGATTTTACAGCAATGTTTACTGATGAAGAGTAAATGGTGAACTGAA 3565

QY 178 TATACCAACAGCCATTCAGGACTTCGGTGTCTATCTACCGTTTATTCGAAAGAAAG 237
Db 3566 AACACACACAGCCATTTATTCAGCATAGTTCGGCATTA-----TTAGCAGCGTAAAA 3619

QY 238 GGTATCAGCTGATATGTTGCTGTTGTTCTCTTGGAGATATCTCTGCTTGGTGCA 297
Db 3620 AATTGAATCCTGATTTTACTATGGGCAATGTTAGTGGAATATTAAGTTTATGCA 3679

QY 298 AGCGGCGCTTGATTTTGAAGATCGGTTGCTTGGTATGCTGAGTAAAGCTGAGCTATATG 357
Db 3680 GCTGAGCTATTATCTTGAAGATGAGTAAATTTGTTAGAAAACGTTGTCATTAATG 3739

QY 358 GAAGAGCGGCTCCTGCTGACTCTGCAAGATGGTAGAGTTCTCAATACGCCAGTAGAG 417
Db 3740 GGCAGAGCATTCCTACTGCTGTTAGGAGCATGGCTGCAGTATGGGATTTAGATTTGAT 3799

QY 418 GTCAATTGAAGAGCCCTGTCA---AAAAGCTTCTGAACITGGAGTGGTTACTCCAGCCAA 474
Db 3800 AAAAGTCGATGAAATTTTGAATGATCATCTATCTGATGACAAATTAATGAACAGCAAC 3859

QY 475 TATAACACACCTGACAAATCTCATCTGAGAGAGTGGTTGAGTTGATCGAGCGGTT 534
Db 3860 ATTAATTGCCAGGTCAAATTTGTTTTCAGGTCAAAAAGCTTTAATTTGATGAGCTAGTA 3919

QY 535 GAACCTTTTCAAGAGAGCAGTGCCTGCAAAAGCTTGTATTCCTTTAAGGTGTCAGGTCCCTTT 594
Db 3920 GAAAAGGTAAATCATTAGTGTGCAAAAGCTGTCGCTTTAGCAGTATCTGACCAATTC 3979

QY 595 CACACCGCTCTCCTTGAGCCTCTAGCAGAAACCTAGCTGAAACTCTTAGCTCAGTAAAGT 654
Db 3980 CATTCATGCTAATGAAAGTGAATGAAGAAAGATTTTCAAGTTACATTAATCAATTTGAA 4039

QY 655 TTTTCAGATTTTACTTGTCCCTTAGTCGCAATACAGAGCT--GCTGTGATGCAAAA 711
Db 4040 TGGCGTATGCTAAGTTTCTCTAGTTTCAAAATGTAATGCGCAAGGTGAACCTGACAAA 4099

QY 712 GAGGACATTTGCTCAGCTCTTGACGCTCAGGTCAGAGAACCGGTTTCGTTTCTATGAAAGT 771
```

```

Db 4100 GAAGTAATTAATCTAATATGTTCAAGCAATTAATTAATCACCAGTACAAATTAATCA 4159
QY 772 ATTGGGCTCATGCAAGAGCAGGATAGCAACTTTATCGAGATTGGACCGGGGAAAGTC 831
Db 4160 ACAGAAATGGCTAATAGACCAAGGTTGATCAATTTATTTGAATTTGCTCTGGAAGATT 4219
QY 832 TTGTCAAGTTTGTAAAAAAATGATCAAACTGCTCACTTAGCTCATGTGGAAGATCAA 891
Db 4220 TTATCTGGCTTAATTAATAAAATAATAGAGATGTTAAGTTAAACATCAATTAACAACTTA 4279
QY 892 GCGAGTTTAGTAGCACCTTTTAGAAA 917
Db 4280 GAAGATGTGAAGGATGGAATGAAA 4305

RESULT 15
US-10-331-061-77
; Sequence 77, Application US/10331061
; Publication No. US20030101486A1
; GENERAL INFORMATION:
; APPLICANT: Lassner, Michael
; APPLICANT: Metz, James G
; APPLICANT: Facciotti, Daniel
; TITLE OF INVENTION: SCHIZOCHYTRIUM PKX GENES
; FILE REFERENCE: CGNE.131.02US
; CURRENT APPLICATION NUMBER: US/10/331,061
; PRIORITY FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/231,899
; PRIOR FILING DATE: 1999-01-14
; PRIOR APPLICATION NUMBER: 60/048,650
; PRIOR FILING DATE: 1997-06-04
; PRIOR APPLICATION NUMBER: 09/090,793
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 77
; LENGTH: 7959
; TYPE: DNA
; ORGANISM: Vibrio marinus
; US-10-331-061-77

Query Match      7.5%; Score 69.2; DB 9; Length 7959;
Best Local Similarity 46.8%; Pred. No. 8.7e-11;
Matches 332; Conservative 0; Mismatches 363; Indels 15; Gaps 3;

QY 152 AAGAGACAACCTCAATCAGACCGCTATACGCAACAGCCATCTAGCGACTTCGGTTG 211
Db 1991 AAGAGAGCAATTAAGTTTAAACGCAACATGCGCAACAGGATTTGATTTGAGTGTG 2050

QY 212 CTATCTACCGTTTATTGCAAGAAAGGGCTATCAGCCTGATATGTTGCTGTTGTTCTC 271
Db 2051 GTCTGTTCAAAACGTTTAAAGCAAGCAGGTTTAAAGCTGATTTTGTGCGGTCATAGTT 2110

QY 272 TTGGAGATACTCTGCGCTTGGTGGAGCGGCGCTTGGATTTTGAAGATCGGTTGCT 331
Db 2111 TCGGTGAGTTAAACCGCATTTATGGGCTGCGGATGTTTGAAGCAAGCAATTAATGATG 2170

QY 332 TGCTAGTAGCGCTGAGCGCTATATGGAAGAACGGCTCTCTGCTGACTCTGCGCAAGATGG 391
Db 2171 TAGCGGTAGTCTGTTGCAAGCAATGCTGCGCCAGCAACAGATTTTGTAT-----G 2224

QY 392 TAGCAGTTTCTAATAGCCAGTAGAGTCAATTTGAAGAACCTGTCAAAAGCTTCTGAAC 451
Db 2225 CAGGTAAAGATGGCGCTGTTGTTGTTGATCCAAAGCAAGTCGCTGTGATCATTTGATACCC 2284

QY 452 TTGGAGTGGTTACTCCAGCAACTATAACACACCTCTCAAAATCGTCATTTGCTGGAGAAG 511
Db 2285 TTGATGATGCTCTATTGTTACTTCACTCGAATACCAAGTTGTTATTTGTTGTTACTA 2344

QY 512 TGTTTCAGTTGATCGAGCGGTTGAATTTTTCAGAAAGCAGGTGCAACCGCTTGAATTC 571
Db 2345 CGGAGCAGGTTGCTGTAGCGGTTTACAACTTAGGTAATGCTGTTTCAAAA---GTTGTC 2401
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 11, 2003, 17:55:46 ; Search time 1424 Seconds
(without alignments)
10474.750 Million cell updates/sec

Title: US-09-308-397-1

Perfect score: 921

Sequence: 1 atgactaaacagcctttt.....tagcacttttagaaaaatag 921

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba:*

2: em_esthm:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hcc:*

9: gb_estl:*

10: gb_est2:*

11: gb_hcc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gssprt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	95.6	10.4	1237	17	AF029428
C 2	87.6	9.5	511	14	BQ18656
C 3	82.2	8.9	544	9	AU069821
C 4	53.2	5.8	427	17	AZ049336
C 5	53	5.8	783	14	BQ96704
6	52.6	5.7	491	12	BF512296

7	44	4.8	744	12	BG597049
8	42.6	4.6	414	13	BJ372892
C 9	42.6	4.8	501	13	BM277916
C 10	42.6	4.6	556	13	BM277853
C 11	42.4	4.6	373	12	BE919981
C 12	42.4	4.6	435	12	BF188564
C 13	42.4	4.6	523	12	BF054427
C 14	42.4	4.6	650	10	BE427401
C 15	42.4	4.6	808	11	AK020696
C 16	41.8	4.5	450	13	BI130047
C 17	41.8	4.5	513	13	BI128381
C 18	41.8	4.5	524	9	AI162058
C 19	41.8	4.5	531	9	AI161931
C 20	41.8	4.5	711	14	BQ246282
C 21	41.2	4.5	684	14	C25558
C 22	41.2	4.5	714	9	AU060245
C 23	41	4.5	469	12	BF050033
C 24	41	4.5	546	12	BF169335
C 25	41	4.5	575	13	BI500434
C 26	40.8	4.4	484	10	BE605118
C 27	40.8	4.4	534	13	BJ236347
C 28	40.8	4.4	658	9	AI162555
C 29	40.8	4.4	658	12	BG594410
C 30	40.6	4.4	500	13	BM278805
C 31	40.6	4.4	672	9	AU033776
C 32	40.2	4.4	405	13	BI132128
C 33	40.2	4.4	425	13	BI129534
C 34	40.2	4.4	462	13	BI129308
C 35	40.2	4.4	470	10	BE500556
C 36	40.2	4.4	474	10	BE590504
C 37	40.2	4.4	522	13	BJ232102
C 38	40.2	4.4	536	13	BJ236100
C 39	40.2	4.4	555	10	BE607044
C 40	40.2	4.4	556	13	BJ236365
C 41	40.2	4.4	557	10	BE637457
C 42	40.2	4.4	562	10	BE591088
C 43	40.2	4.4	569	10	BE435529
C 44	40.2	4.4	574	14	BQ245331
C 45	40.2	4.4	606	10	BE422581

ALIGNMENTS

RESULT 1
AF029428/c
LOCUS AF029428 Salmonella typhimurium LT2, linear GSS 29-AUG-2000
DEFINITION typhimurium genomic clone 1257-T7, DNA sequence.
ACCESSION AF029428
VERSION AF029428.1 GI:2570958
KEYWORDS GSS.
SOURCE Salmonella typhimurium.
ORGANISM Salmonella typhimurium.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE 1 (bases 1 to 1237)
Wong,R.M.-Y., Wong,K.K., Benson,N.R. and McClelland,M.
TITLE Sample sequencing of a Salmonella typhimurium LT2 lambda library:
comparison to the Escherichia coli K12 genome
JOURNAL FEMS Microbiol. Lett. 173 (2), 411-423 (1999)
MEDLINE 99243757
COMMENT Contact: McClelland M
Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA
Email: mclelland@lifsci.sdsu.edu
Class: shotgun.
Location/Qualifiers
1. 1237
/organism="Salmonella typhimurium"
/strain="LT2"
/db_xref="taxon:602"

```

/clone="1257-T7"
/clone.lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor
sequencer"
BASE COUNT      293 a      345 c      304 g      281 t      14 others
ORIGIN
Query Match      10.4%; Score 95.6; DB 17; Length 1237;
Best Local Similarity 54.7%; Pred. No. 1.1e-17;
Matches 217; Conservative 7; Mismatches 167; Indels 6; Gaps 2;

QY      1  ATGACTAAACAGACGCTTTTATTGTCGTGGTCAAGGTGCCAGGTATCTAGGAGATGGACGG 60
Db      398  ATGACGCAATTGGATTGTTGCTCCCGCGTCAGGGTTCTCAGAGCGTTGGGATGTGGCC 339

QY      61  GATTTCTATGATCAGTATCCGATTGTCAAGAAACGATTGATCGAGCGAGTCAGGTGCTC 120
Db      338  GAGATGGCGGCAAAATACCCTATCGTAGAAGAAACGTTTGCTGAAGCTTCTCGGCTCTG 279

QY      121  GCTTATGATTACGTATTATCTCATCGATACCGA---AGAAGACAAACTCAATCAGACCCGC 177
Db      278  GGATATGATCTGTGGCGGCTCACCAGCAAGTCCAGCGGAAGAACTGAATAAAACCTTG 219

QY      178  TATACGCACACAGACCAATCTAGCGACATCCGGTGTCTATCTACCGTTTATGCAAGAAAG 237
Db      218  CAGACGACGCGGCTATTAAACGCTTCCGNTSCGNTWTGGCGCTTTGGCAGCAGCAG 159

QY      238  GGCATATCA---GCCGTATAGTTGCTGGTTTCTCTCTTTGGAGAAATATCTTGTCCTTGGTG 294
Db      158  GCGCGTNAAAATGCGCTGCGTAAATGGCAGGTACAGAGCTGGCGGAATATTTCGCGCTGTT 99

QY      295  GCAAGCGCGGCTTGGATTTTGAAGATGCGGTTGCCCTTGCTAGCTAAGCGTGGAGCCTAT 354
Db      98  TGMGCTGCGGCTATCAACTTTGCTGATCCGTTCTGCTGTGGAATGCGCGGTAATTC 39

QY      355  ATGGAAGAACGGGTCTCTGCTGACTCTGGCAAGATGG 391

Db      38  ATGCAGGAAGCAACCGGAGAACTGGCGCATGTGG 2

```

```

RESULT 2
EQ18656/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
  ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

B0818656          511 bp      mRNA      linear      EST 01-AUG-2002
1030072D02.y1 C. reinhardtii CC-1690, Deflagellation (normalized),
Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.

B0818656
B0818656.1 GI:22068897
EST.
Chlamydomonas reinhardtii.
Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.

1 (bases 1 to 511)
Grossman, A., Chang, C.-W., Davies, J., Harris, E., Hauser, C., Lefebvre
P., McDermott, J. P., Shreeve, J., Silflow, C. and Stern, B.
Analyses of the Chlamydomonas reinhardtii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Vascular Plants. Project: 1030
Unpublished (2002)
Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
Email: chauser@duke.edu.

1. .511
Location/Qualifiers
/organism="Chlamydomonas reinhardtii"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
/clone_lib="C. reinhardtii CC-1690, Deflagellation
(normalized), Lambda Zap II"
FEATURES
source

```

/notes=Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; DeLagallation library, constructed by John Davies and Jeffrey McDermott, combines cDNAs from CC-1690 cells which had been re-synthesized by pH shock. polyA mRNA was min after being dephosphorylated by pH shock. polyA mRNA was purified from each sample, pooled and cDNA synthesized. The cDNA was directionally cloned into lambda Zap II (Stratagene) in the EcoRI (5') and XhoI (3') sites. pBluescript II SK- plasmids were excised from the lambda Zap clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described in Bonaldo et al., (1996) Genome Research 6: 791-806."

BASE COUNT	107 a	148 c	126 g	130 t
ORIGIN				
Query Match	9.5%	Score 87.6;	DB 14;	Length 511;
Best Local Similarity	51.3%;	Pred. No. 1.7e-15;		
Matches	229;	Conservative	0;	Mismatches 214; Indels 3; Gaps 1;
QY	421	ATTGAAGAGCCCTGTC	CAAAAAGCTTCGAACT	TGGAGTGGTTACTCAGGCCACTATAAC
Db	511	ATTGCGAAAGCGGTG	GAAGAAGTCGAGAAGT	CGGTGCTGTTCTCCGGTAAACTTTAAC
QY	481	ACACCTGCACAAATCG	TATTCGTCGAGAAGT	GGTTGTCAGTTGATCGAGCGGTTGAACCTT
Db	451	TCTCCGGACAGCTGG	TATTGCGCGTCATAA	GAGCGGTGAGCGTGTGCGCGTGC
QY	541	TTGCAAGAGCAGGTG	CCCAAAACGCTTGAT	TCTTAAAGTGTACAGTCCCTTTCACACC
Db	391	TGTAAGAAGCGCGG	CGCAAAACGCGCGT	TCCGTTTACCAGTGAAGTACCGTCTCACTGT
QY	601	GCCTCTCTTGAGCCT	GTAGCCAGAAACTAG	CTGAAACTCTAGCTCAGGTAACTTTTCA
Db	331	GCCTGATGAACAC	GACGCGCAAACTG	GCGAGTAGAATTAGCGAAATCACCTTTAAC
QY	661	GATTTTACTTTGTC	CCCTAGTCGGCAAT	CAGAAAG---CTGCTGTGATGCAAAAGAGGAC
Db	271	GCACCAACACAGT	CTCCTGTGTGAATA	TAACGTTGATGCAAAATCGAAACCAATGCGTATGCC
QY	718	ATTGCTCAGCTCTT	GACGGTGTAGGTC	CAAGGACCGTTGCTTTCTATGAAGCTATTGGG
Db	211	ATCCGTGACCACTG	TAGTCACTTGTAT	AACCCGGTTCAGTGGAGGAGTCTGTTGAG
QY	778	GTCATGCAAGAAGC	AGGCATAGAAGCA	ACTTTATCGAGATTGGACCGGGAAAGTCTTGTCA
Db	151	TACATGCGACGCG	CAAGCGGTAGAAC	CTCTATGAAGTGGCCCGGCAAGGTCTACT
QY	838	GGTTTTGTTAAAAA	ATTGNTCAAC	863
Db	91	GGCGTGACGAAAC	GCATTTGTCGAC	66

RESULT 3
AU069821/c
LOCUS
DEFINITION
Rice panicle at flowering stage Oryza sativa (japonica cultivar-group) cDNA clone E3726_7A, mRNA sequence.
AU069821
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGNISM
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 544)
Sasaki, T. and Yamamoto, K.
Rice cDNA from panicle at flowering stage
Unpublished (1996)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan

Tel: 81-298-38-7441
 Fax: 81-298-38-7468
 Email: tsasaki@affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
 PROJECT = RGP

FEATURES
 source
 Location/Qualifiers
 1. 544
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="E3726 7A"
 /clone_lib="Rice panicle at flowering stage"
 /dev_stage="flowering stage"
 /note="Organ: panicle; Rice cDNA from panicle at flowering stage"

BASE COUNT 118 a 161 c 126 g 133 t 6 others
 ORIGIN

Query Match 8.9%; Score 82.2; DB 9; Length 544;
 Best Local Similarity 51.4%; Pred. No. 8.3e-14;
 Matches 280; Conservative 0; Mismatches 259; Indels 6; Gaps 4;

QY 316 GAAGATCGGTTGCTTGGTAGCTAAGCGTGGAGCTATATGGAAGAGCGCTCCTGCT 375
 DB 544 GNTGATCGGTCGCTGCTGGTAGTGCNCGCAAGTTCAAGCAAGCCGTACCGAA 485
 QY 376 GACTCTGCAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGTC-ATTGAAGAGCTG 434
 DB 484 GCGAGGCGGTATGCGGCAATCATCGTCTGGAATGATCGTCTATTGGAAGCGTG 425
 QY 435 TCAAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGCCCACTATAACACCTGCACAAAT 494
 DB 424 TNAAGAAGCTGCAGAAGGTGAGTGGTTCTCCGGTAAACTTTAACTCTCC-GGACAGGT 366
 QY 495 GGTATTGCTGAGAGAGTGGTTGCAAGTGTATCGAGCGGTGAAGCTTTTGAAGAGCAGG 554
 DB 365 GGTATTGCTGAGAGAGTGGTTGCAAGTGTATCGAGCGGTGAAGCTTTTGAAGAGCAGG 306
 QY 555 TGCC-AAAGCTGATCTCTTAAAGTGTAGTCCCTTTACACCGCTCTCTCTGAGC 613
 DB 305 GCGNAAAGCGCGCTGCGGTACAGTGAGCGGTACCGTCTCACTGTGCGCTGATGAAC 246
 QY 614 CTGTAGCGCAGAACTAGCTGAACTCTAGCTCAGGTAAAGTTTTCAGATTTTACTTTGTC 673
 DB 245 CAGCAGCGCAGAACTAGCTGAGTAAATAGCGAAATCACTTTAAGCACCACAGTTC 186
 QY 674 CCTAGTGGCAATACAGAAG---CTGCTGTATGCAAAAGAGCAATGCTCAGCTCT 730
 DB 185 CTGTTGTGAATAACGTTGATGTGAATCGGAAACCAATGGTGATCGCATCGTGACGCAC 126
 QY 731 TGAGCGCTCAGTCAAGGAACCGTTCTGTTTCTATGAAGTATTGGGTCTATGCAAG 790
 DB 125 TGGTACGTGATTTGATATACCGGTTTCTAGTGACAGAGTCTGTTGAGTACATGCGCGC 66
 QY 791 CAGGCATTAAGCAACTTTATCGAGATTGACCGGGGAAAGTCTTCTCAGGTTTCTTAAAA 850
 DB 65 AAGCGGTAGAACATCTCTATGAAGTCGCGCGGCAAGTCTTACTGCGCTGACGAAC 6
 QY 851 AAATT 855
 DB 5 GCATT 1

RESULT 4
 LOCUS AZ049336/c
 DEFINITION GS8Br0899 Sheared genomic library Brucella melitensis biovar Abortus genomic clone Bc71, DNA sequence.
 ACCESSION AZ049336
 VERSION AZ049336.1 GI:7273251
 KEYWORDS GSS.
 SOURCE Brucella melitensis biovar Abortus.
 ORGANISM Brucella melitensis biovar Abortus
 Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

Brucellaceae: Brucella.
 1 (bases 1 to 427)
 Sanchez,D.O., Zandomeni,R.O., Cravero,S., Verdun,R.E., Pierrou,E., Faccio,P., Diaz,G., Lanzavecchia,S., Aguerro,F., Frasch,A.C.C., Andersson,S.G.E., Rosetti,O.L., Grau,O. and Ugalde,R.A.
 Gene discovery through genomic sequencing of Brucella abortus Infect. Immun. 69 (2), 865-868 (2001)
 21101034
 Contact: Zandomeni, R.
 Centro de investigacion en Ciencias Agropecuarias (CICIA)
 Instituto Nacional de Tecnologia Agropecuaria (INTA) C.C. 25 (1712)
 Castelar. Buenos Aires, Argentina
 Tel: 5411-4621-3316/1683
 Fax: 5411-4481-1316
 Email: zandomeni@inta.gov.ar
 Class: shotgun.

FEATURES
 source
 Location/Qualifiers
 1. 427
 /organism="Brucella melitensis biovar Abortus"
 /strain="S-2308"
 /db_xref="taxon:235"
 /clone="Bc71"
 /clone_lib="Sheared genomic library"
 /note="Vector: pBluescript SK(-) (STRATAGENE); Genomic DNA was mechanically sheared, blunt ended, and size-fractionated by agarose gel electrophoresis. Fragments between 1.5-3 Kb were recovered and ligated to the pBluescript SK (-) vector."
 BASE COUNT 86 a 145 c 120 g 74 t 2 others
 ORIGIN

Query Match 5.8%; Score 53.2; DB 17; Length 427;
 Best Local Similarity 48.1%; Pred. No. 6e-05;
 Matches 151; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 256 GTTCGCTGTTCTCTTGGAGAACTCTGCTGGTGGCAAGCGGCGCTTGGATTT 315
 DB 328 GTGCGCGGTATCGTCTGGCAATATCGGCTCTGTGCGCTGCTGGACGTTTTCATT 269
 QY 316 GAAGATGCGGTGCTTGGTAGCTAAGCGTGGAGCTATATGAAGAAGCGGCTCCTGCT 375
 DB 268 GCGCACGCGCGTCTCTTGGCATTCGCGCAATGCCATCAGAGCGGTTCGGTC 209
 QY 376 GACTCTGGCAAGATGTAGCAGTTCTCAATACGCCAGTAGAGTCAATTGAAGAAGCTGT 435
 DB 208 GCGCAAGCGGCTATGCGCGCATCATCGCTCTGGAACACGCGCATGTGGAGCTGTCTGC 149
 QY 436 CAAAAGCTTCTGAACCTTGGAGTGGTTACTCCAGCAACTATAACACACTGCACAAATC 495
 DB 148 ATGGAAGCAAGAGCTTCGCGCCGTCAGATCGGCATGCAATGGCGCGCCAGCTT 89
 QY 496 GTCAATGCTGGAGAGTGGTTTCAGTTGATCGAGCGGTGAACTTTTGAAGAAGCAGGT 555
 DB 88 GTCATATCGGCGCCAGGCGCGCTTGAATCGCGCGCAAGCTTGGCTCTCGAAAAGGCT 29
 QY 556 GCGAAGCGTTTGCAT 569
 DB 28 GCAAGCGCGCCAT 15

RESULT 5
 LOCUS BQ796704/c
 DEFINITION EST 5642 Ripening Grape berries Lambda zap II library Vitis vinifera cDNA clone R1051B09 3', mRNA sequence.
 ACCESSION BQ796704
 VERSION BQ796704.1 GI:22011670
 KEYWORDS EST.
 SOURCE Vitis vinifera.
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae; Vitis.

```

REFERENCE 1 (bases 1 to 783)
AUTHORS Abbal,P., Agasse,A., Ageorges,A., Atanassova,R., Barrieu,F.,
        Couture,C., Dedaldechamp,F., Delrot,S., Glissant,D., Grimplet,J.,
        Hamdi,S., Romieu,C. and Terrier,N.
TITLE Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
or seeds) at Various Developmental Stages
JOURNAL Unpublished (2002)
COMMENT Contact: Romieu C.
        Unite de Recherche des Produits de la Vigne
        Institut National de la Recherche Agronomique
        2, place Viala, 34 060 Montpellier Cedex 01, France
        Tel: 00-33-(0)4-99-61-28-62
        Fax: 00-33-(0)4-99-61-28-57
        Email: romieu@ensam.inra.fr
        Seq primer: I7.

FEATURES             Location/Qualifiers
     source           1..783
                     /organism="Vitis vinifera"
                     /cultivar="Shiraz"
                     /db_xref="taxon:29760"
                     /clone="PT051B09"
                     /clone_lib="Ripening Grape berries Lambda Zap II Library"
                     /dev_stage="ripening stage"
                     /note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco RI
                     ; Site 2: XhoI; Oriented library, construction described
                     in Generation of ESTs from Grape Berry (skin, pulp or
                     seeds) at various developmental stages by Terrier,N.,
                     Ageorges,A., Abbal,P., Romieu,C. in J. Plant Physiol. 158
                     (12): 1575-83 2001"

BASE COUNT 220 a 181 c 145 g 237 t
ORIGIN
Query Match          5.8%; Score 53; DB 14; Length 783;
Best Local Similarity 45.8%; Pred. No. 0.0001;
Matches 222; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 425 AAGAGCGCTCTAAAAGCTTCTGAAGTGGAGTGTACTCCAGCACTAATACAC 484
      |||||
Db 783 ATCAGAGCAATGAAGAAGTGTGATGAATGAATGAATGCGAAATTCCTATGCC 724
      |||||
QY 485 CTGCACAAATCGTCATTCTGCGAGAGAGTGTTCAGTTGATCGAGCGTTGAACCTTTGC 544
      |||||
Db 723 CTGGGATTTATGAGTCTCTGGAGGAGTGAAGAGTGGAGCGTTGAAGCTAAGGCTA 664
      |||||
QY 545 AAGAGCAGGTGCCAAGCGTTGATTCCTTTAAGTGTGAGTCCCTTTTCACCGCTC 604
      |||||
Db 663 AGTCATTCGAAGGCCGAATGACGGTGGCTTTCAGTGGCTGTGCTTTCCACACTCATT 604
      |||||
QY 605 TCCTTGAGCTGTGACGCAAACTAGCTGAACTCTAGCTAGTAACTTTTCAGATT 664
      |||||
Db 603 TCATGGAACTGTGCTCTCAAGATTGGAAGCCACATGCGAGCTACGGAATCAGAACTC 544
      |||||
QY 665 TTACTTGTCCCTAGTGGCAATACAGAAGCTCTGTGATGCAAAAGAGGAC---ATTG 721
      |||||
Db 543 CAAGAATCCAGTTATATCAAGCTTGATGCACAGCCACATGCAGATCCGACACGATTA 484
      |||||
QY 722 CTGAGCTCTTGACGCTGAGTCAAGAACCCGTTCTGTTCTATGAAGATTTGGGGTCA 781
      |||||
Db 483 AGAAGATATTGGCGCGCTCAAGTGAAGTCTCCCACTTCTATGGGAAACACAGTGAAGCTC 424
      |||||
QY 782 TGAAGAAGCAGGATGAAGCACTTTATGAGATTGGACCGGGGAAGTCTTGTCAGGTT 841
      |||||
Db 423 TCCTNACCAAGAGCTATAAAGAGTTACGAATTGGGGCTGGAAGGTTATCGTGGCA 364
      |||||
QY 842 TTGTTTAAAAATGATCAAACTGCTCACTAGCTCATGTGGAAGATCAAGCGAGTTTATG 901
      |||||
Db 363 TTTTCAAGAGATGGAACAAACTGCCGACTTAGAAGATGTTAGTGTGATTGAGTGGG 304
      |||||
QY 902 TAGCA 906
      |||||
Db 303 AATCA 299
      |||||

```

RESULT 6
 BF512296
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 491)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: NCI-CGAP clone distribution
 information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward
 POLYA-No.

Location/Qualifiers
 1..491
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3069391"
 /clone_lib="NCI CGAP Sub7"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: p773D-pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; NCI CGAP Sub7
 is a subtracted library derived from NCI CGAP Sub6. The
 NCI CGAP Sub7 library had 12 million recombinants. A
 single-stranded DNA preparation of NCI CGAP Sub6 was used
 as a tracer in a subtractive hybridization with a driver
 comprising the IMAGE pool (NCI CGAP Kid3 pool 1 LLAM
 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)
 132378-132391, 145608-145675, 150052-150285);
 NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
 (IMAGE Clones) 132391-132583, 1471368-1472903,
 1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582,
 3851-3854 (IMAGE Clones) 1414920-1417991, 1520904-1522439
); NCI CGAP GC4 pool 1 LLAM 3164-3167, 3716-3720,
 3733-3735 (IMAGE Clones) 1257096-125831, 1469064-1470983
 , 1475592-1476743); NCI CGAP P22 pool 1 LLAM 2457-2459,
 2758-2759, 3062-3068 (IMAGE Clones) 985608-986759,
 1101192-1101959, 1217928-1220615); NCI CGAP C610 pool 1
 LLAM 2644-2653, 2871-2872 (IMAGE Clones) 1057416-1061255
 , 1144584-1145351). (6% of the driver population), plus a
 pool of 3,840 arrayed clones from NCI CGAP Sub1 (IMAGE
 Clones) 2708616-2710535 and NCI CGAP Sub2 (IMAGE
 Clones) 2710536-2712455 (4% of the driver population
), plus a pool of 11,136 clones from NCI CGAP Sub3 (IMAGE
 Clones) 2712456-2723591 (10% of the driver population),
 plus a pool of 5,472 clones from NCI CGAP Sub4 (IMAGE
 Clones) 2723592-2729326 (40% of the driver population),
 plus a pool of 4032 clones from NCI CGAP Sub6 (IMAGE
 Clones) 2728969-2733190 (40% of the driver population).
 Subtraction was performed as previously described (Bonaldo
 , Lennon & Soares (1996): Normalization and Subtraction:
 Two Approaches to Facilitate Gene Discovery. Genome
 Research 6, 791-806.
 TAG SEQ=None found"

BASE COUNT 90 a 152 c 167 g 82 t
 ORIGIN

Query Match 5.7%; Score 52.6; DB 12; Length 491;
 Best Local Similarity 50.6%; Pred. No. 0.0001;
 Matches 127; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

```

QY 452 TTGAGTGTGTTACTCCAGCCAACTATAACACACCTGCAAAATCGTCATTGCTGAGAAG 511
Db 124 TCGAGGTGTGGAAGCGGTGAACCTCAACGACCCCGGCGAGACCGTGATCGCGGGAGCA 183
QY 512 TGTGTCAGTTGATCGAGCGGTGAACCTTTTCAGAGACGAGTGCCTCCAAAGCTTGATTC 571
Db 184 AGCGCGGTGCAACAGGCGCTGTGAGGTGCTCAAGGCCAACCGTGCAGAGCGCGCTTGC 243
QY 572 CTCTTAAGGTGTCAGTGCCTTTTCACCGCTCTCTTGAGCTGCTAGCCAGAACTAG 631
Db 244 CATTGCGGTGTCGCGACCGTTCCTCATTCAGCTCATGAGCGCGTCCGAGCGTCTGA 303
QY 632 CTGAAGACTCTAGTCAAGTAACTTTTTCAGATTTTCTGTCCTTCCCTAGTCGCAATACAG 691
Db 304 AAGAAAGCTCCCAATACGGTGTGTTTCCAGCGCAGATCCCGGTGCTCAACACATCG 363
QY 692 AAGCTGCTGTG 702
Db 364 ATGTGGCGGTG 374

RESULT 7
LOCUS BG597049 744 bp mRNA linear EST 12-APR-2001
DEFINITION EST495727 cSTS Solanum tuberosum cDNA clone cSTS16111 5' sequence, mRNA sequence.
ACCESSION BG597049
VERSION BG597049.1 GI:13615189
KEYWORDS EST.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 744)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho J., Chiemiango,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13P-R.

FEATURES
source
location/Qualifiers
1..744
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS16111"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stages="12-14 weeks post harvest"
/lab_host="SOUR"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 25C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."
BASE COUNT 188 a 149 c 185 g 222 t
ORIGIN

Query Match 4.8%; Score 44; DB 12; Length 744;
Best Local Similarity 54.3%; Pred. No. 0.059;
Matches 89; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 229 CAAGAAAGGCGCTATCACCTGATGTTGCTGCTGTTGCTCTCTTGGAGATACTCTGCC 288
Db 485 CAACAAATAATGATTCGGTGTGATGTACATGTGCTGAGTTTGGCGGAATACACTGCT 544
QY 289 TTGTGGCAAGCGGCGCTTGGAATTTTCAGATCGGCTTGCTTGGTGTAGCTAAGCGTGA 348

```

```

Db 545 CTGGCATTTACAGAGCTTTCAGTTTGGATGGCTTAAGCTTGAAGCTTAGAGGA 604
QY 349 GCCTATATGAAAGAGCGGCTCCTGCTGACTCTGCGAAGATGGT 392
Db 605 GAAGCTATGCAAGATGCTGCTGATGCTGCACAAAAGTGCAATGGT 648

RESULT 8
LOCUS BJ372892 414 bp mRNA linear EST 08-MAR-2002
DEFINITION BJ372892 Dictyostelium discoideum cDNA library, CF Dictyostelium discoideum cDNA clone ddc14f21 3', mRNA sequence.
ACCESSION BJ372892
VERSION BJ372892.1 GI:19282275
KEYWORDS EST.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
REFERENCE 1 (bases 1 to 414)
AUTHORS Urushihara,H., Tanaka,Y., Kohata,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoideum at the culmination stage
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
location/Qualifiers
1..414
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc14f21"
/clone_lib="Dictyostelium discoideum cDNA library, CF"
/sex="mat A"
/dev_stage="Culmination stage"
BASE COUNT 98 a 42 c 152 g 116 t
ORIGIN

Query Match 4.6%; Score 42.6; DB 13; Length 414;
Best Local Similarity 48.1%; Pred. No. 0.11;
Matches 117; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 314 TTGAAGATCGGTTGCCTTGGTAGCTAAGCGTGGAGCCTATATGGAAGAAGCGGCTCCTG 373
Db 50 TTGTAGTTGCAGCTGTGGTTGTAGCTATGTTGCAGCTGCTGTTGTAGCTGTTGTGG 109
QY 374 CTGACTCTGGCAAGATGGTAGCAGTTCTCAATACGCCAGTAGAGGTCAATTGAAGAGCT 433
Db 110 CTGTGTTGTAGCTGTGGTTGGGTTGTAGTTGCAGTTGTAGTTGCAGCTGTAGTTGCA 169
QY 434 GTCAAAAACCTTCTGAACCTTGAGTGGTTACTCCAGCCAACTATACACACCTGCACAA 493
Db 170 TTGTAGTTCCAGCTGTAGTTGCTGCAGATGTCAGATGAAGATGATGCAGATGAAGTG 229
QY 494 TCCTCATTCTCGAGAAGTGGTTGCAGTTGTATCGAGCGGTTTCAACTTTTTCGAAGAGCAG 553
Db 230 GANCAGAGCTTGATGATGCAGATGAAGATGGGGCAGAGCTTGATGATGCAGATGAANATG 289
QY 554 GTG 556
Db 290 GGG 292

RESULT 9
LOCUS BM277916/c 501 bp mRNA linear EST 20-DEC-2001
DEFINITION BM277916 As_tgz_50D12_SKPL Ascaris suum adult male testis germinal zone from Alan Scott Ascaris suum cDNA clone As_tgz_50D12 5', mRNA sequence.

```



```

ACCESSION      BM277916.1  GI:17971174
VERSION        EST. roundworm.
KEYWORDS       Ascaris suum
SOURCE         Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
ORGANISM       1 (bases 1 to 501)
REFERENCE      Blaxter M.L., Parkinson J., Whitton, C., Daub, J., Guiliano, D., Hall
AUTHORS        N., Quayle, M. and Barrell, B.
TITLE          Edinburgh University/Sanger Centre Nematode EST Project
JOURNAL        Unpublished (2000)
COMMENT        Contact: Blaxter ML
                Institute of Cell, Animal and Population Biology
                University of Edinburgh
                Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
                3JT, UK.
                Tel: +44 131 650 6760
                Fax: +44 131 670 5450
                Email: mark.blaxter@ed.ac.uk
                The library was prepared by Michelle Lizotte-Waniewski for Alan
                Scott, Johns Hopkins University Medical School, Baltimore.
                Sequencing was performed by the Pathogen Sequencing Unit, Sanger
                Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).
                PCR Primers
                FORWARD: T3
                BACKWARD: T7PL
                Plate: 50 row: D column: 12
                Seq primer: SKPL
                High quality sequence stop: 446.
                Location/Qualifiers
                1..501
                /organism="Ascaris suum"
                /db_xref="taxon:6253"
                /clone="As tgz 50012"
                /clone_lib="Ascaris suum adult male testis germinal zone
                from Alan Scott"
                /sex="Male"
                /dev_stage="Adult"
                /note="Vector: Lambda Zap II; Site 1: EcoRI; Site 2: XhoI;
                Library was made from dissected testis germinal zone from
                adult male Ascaris suum collected from abattoirs.
                Constructed by Michelle Lizotte-Waniewski for Alan Scott,
                Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT    187 a 143 c 58 g 113 t
ORIGIN
Query Match      4.6%; Score 42.6; DB 13; Length 501;
Best Local Similarity 48.2%; Pred. No. 0.12;
Matches 120; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 308 TGGATTTCAGATCGGTTGCTGGTAGCTAAGCGTGAGCCCTATATGGAAGACGG 367
Db 321 TTGAAGTTCTGCTGCTGTTGAAGTTGCTGTTGCTAAGGTTGTTGAGGTTATTGTTGAGG 262
QY 368 CTCCTGCTGACTCTGCAAGATGGTAGCAGTTCTCAATACGCCAGTACAGGTCATTGAAG 427
Db 261 TTGCTGAGGTTGTTGTGAGTAGTTAGGTTGTTATTGAGTTGTTGAGTTGTTCAAG 202
QY 428 AAGCCGTGCAAAAGCTTCTGAACCTTGAGTGGTGGTACTCCAGCCAACTATACACACCTG 487
Db 201 TTGTAGCTGAGGTTGCTCAGAGAGTTGCTGAGGTTGTTATTGAAGTTATTTCAGAGAGTTG 142
QY 488 CACAAATCGTCAATTGCTGGAGAAGTGGTTGCAGTTGATCGAGCGGTTGCAACTTTTGAAG 547
Db 141 CTGAGGTTGTTATTGAAGTTGCTGTTGTTGAAGTTGCTGTTGAGTTCTGTTGAGTTCTGTTGAAA 82
QY 548 AAGCAGGTG 556
Db 81 ATGCTATTG 73
RESULT 10

```

```

BM277853/C
LOCUS          BM277853          556 bp      mRNA      linear      EST 20-DEC-2001
DEFINITION    As tgz_49E05 SKPL Ascaris suum adult male testis germinal zone from
              Alan Scott Ascaris suum cDNA clone As tgz_49E05 5', mRNA sequence.
ACCESSION     BM277853
VERSION       BM277853.1  GI:17971111
KEYWORDS      EST.
SOURCE        pig roundworm.
ORGANISM      Ascaris suum
REFERENCE      Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea
              ; Ascarididae; Ascaris.
AUTHORS        1 (bases 1 to 556)
              N., Quayle, M. and Barrell, B.
              Blaxter M.L., Parkinson J., Whitton, C., Daub, J., Guiliano, D., Hall
              Edinburgh University/Sanger Centre Nematode EST Project
              Unpublished (2000)
TITLE          Contact: Blaxter ML
JOURNAL        Institute of Cell, Animal and Population Biology
COMMENT        University of Edinburgh
              Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
              3JT, UK.
              Tel: +44 131 650 6760
              Fax: +44 131 670 5450
              Email: mark.blaxter@ed.ac.uk
              The library was prepared by Michelle Lizotte-Waniewski for Alan
              Scott, Johns Hopkins University Medical School, Baltimore.
              Sequencing was performed by the Pathogen Sequencing Unit, Sanger
              Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell). The
              sequence contained a PolyA tail (trimmed)
              PCR Primers
              FORWARD: T3
              BACKWARD: T7PL
              Plate: 49 row: E column: 05
              Seq primer: SKPL
              High quality sequence stop: 542.
              Location/Qualifiers
              1..556
              /organism="Ascaris suum"
              /db_xref="taxon:6253"
              /clone="As tgz_49E05"
              /clone_lib="Ascaris suum adult male testis germinal zone
              from Alan Scott"
              /sex="Male"
              /dev_stage="Adult"
              /note="Vector: Lambda Zap II; Site 1: EcoRI; Site 2: XhoI;
              Library was made from dissected testis germinal zone from
              adult male Ascaris suum collected from abattoirs.
              Constructed by Michelle Lizotte-Waniewski for Alan Scott,
              Johns Hopkins University Medical School, Baltimore, MD."
BASE COUNT    202 a 154 c 76 g 124 t
ORIGIN
Query Match      4.6%; Score 42.6; DB 13; Length 556;
Best Local Similarity 48.2%; Pred. No. 0.13;
Matches 120; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
QY 308 TGGATTTCAGATCGGTTGCTGGTAGCTAAGCGTGAGCCCTATATGGAAGACGG 367
Db 371 TTGAAGTTCTGCTGCTGTTGAAGTTGCTGTTGCTAAGGTTGTTGAGGTTATTGTTGAGG 312
QY 368 CTCCTGCTGACTCTGCAAGATGGTAGCAGTTCTCAATACGCCAGTACAGGTCATTGAAG 427
Db 311 TTGCTGAGATTGTTGTTGAGATAGTTGAGTTGTTATTGAGGTTGTTGAGTTGTTCAAG 252
QY 428 AAGCCGTGCAAAAGCTTCTGAACCTTGAGTGGTGGTACTCCAGCCAACTATACACACCTG 487
Db 251 TTGTAGCTGAGGTTGCTCAGAGAGTTGCTGAGGTTGTTATTGAAGTTATTTCAGAGAGTTG 192
QY 488 CACAAATCGTCAATTGCTGGAGAAGTGGTTGCAGTTGATCGAGCGGTTGCAACTTTTGAAG 547
Db 191 CTGAGGTTGTTATTGAAGTTGCTGTTGTTGAAGTTGCTGTTGAGTTGCTGTTGAGTTGTTGAAG 132
QY 548 AAGCAGGTG 556

```

```

Db      131 TTGCTATTG 123
||
||
RESULT 11
BE919981
LOCUS
DEFINITION EST423750 potato leaves and petioles Solanum tuberosum cDNA clone
ACCESSION BE919981
VERSION BE919981
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 373)
Tel: (301)838-0200
Email: http://www.tigr.org/tdb/potato/
For clone information, please contact Research Genetics, Libraries
Division, tel# 1-800-711-6195, email- cdna@resgen.com.
FEATURES
Location/Qualifiers
1..373
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB3E24"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 101 a 60 c 104 g 108 t
ORIGIN
Query Match 4.6%; Score 42.4; DB 12; Length 373;
Best Local Similarity 53.7%; Pred. No. 0.12;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy 229 CAAGAAAGGCTATCAGCTGATATGTTGGTTCCTTCTTGGAGAACTCTGCC 288
Db 164 CAACAAATAATTGATTCGGTTGATGTCACATGTGCTGAGCTTGGCGGAATACACTGCT 223
Qy 289 TTGTTGGCAGCGCGCTTGGATTTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGA 348
Db 224 CTGCATTACAGAGCTTTCAGTTTTCAGATGGGCTTAAGCTGGTTAAGCTTAGAGGA 283
Qy 349 GCCTATATGGAAGAGCGGCTCTGCTGACTCTGGCAAGATGCT 392
Db 284 GAAGCTATGAGGATGCTGCTGATGCTGCAAAAAGTCAATGCT 327
RESULT 12
BF188564
LOCUS
DEFINITION EST444851 potato stolon, Cornell University Solanum tuberosum cDNA
ACCESSION BF188564
VERSION BF188564
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 523)
Tel: (301)838-0200
Email: http://www.tigr.org/tdb/potato/
For clone information, please contact Research Genetics, Libraries
Division, tel# 1-800-711-6195, email- cdna@resgen.com.
FEATURES
Location/Qualifiers
1..523
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="CSTA45G17"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT 119 a 73 c 118 g 125 t
ORIGIN
Query Match 4.6%; Score 42.4; DB 12; Length 435;
Best Local Similarity 53.7%; Pred. No. 0.13;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy 229 CAAGAAAGGCTATCAGCTGATATGTTGGTTCCTTCTTGGAGAACTCTGCC 288
Db 148 CAACAAATAATTGATTCGGTTGATGTCACATGTGCTGAGCTTGGCGGAATACACTGCT 207
Qy 289 TTGTTGGCAGCGCGCTTGGATTTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGA 348
Db 208 CTGGCATTACAGAGCTTTCAGTTTTCAGATGGGCTTAAGCTGGTTAAGCTTAGAGGA 267
Qy 349 GCCTATATGGAAGAGCGGCTCTGCTGACTCTGGCAAGATGCT 392
Db 268 GAAGCTATGAGGATGCTGCTGATGCTGCAAAAAGTCAATGCT 311
RESULT 13
BF054427
LOCUS
DEFINITION EST439657 potato leaves and petioles Solanum tuberosum cDNA clone
ACCESSION BF054427
VERSION BF054427
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 523)
Tel: (301)838-0200
Email: http://www.tigr.org/tdb/potato/
For clone information, please contact Research Genetics, Libraries
Division, tel# 1-800-711-6195, email- cdna@resgen.com.
FEATURES
Location/Qualifiers
1..435
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="CSTA45G17"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
RNA was isolated from developing axillary buds of potato
nodal stem cuttings cultured on medium for the
introduction of tuber formation as described in Bachem et
al. (Plant Journal 1996). Tissue samples were taken of
stages corresponding to growing stolons and the early
stages of tuber formation."
BASE COUNT 119 a 73 c 118 g 125 t
ORIGIN
Query Match 4.6%; Score 42.4; DB 12; Length 435;
Best Local Similarity 53.7%; Pred. No. 0.13;
Matches 88; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
Qy 229 CAAGAAAGGCTATCAGCTGATATGTTGGTTCCTTCTTGGAGAACTCTGCC 288
Db 148 CAACAAATAATTGATTCGGTTGATGTCACATGTGCTGAGCTTGGCGGAATACACTGCT 207
Qy 289 TTGTTGGCAGCGCGCTTGGATTTTGAAGATGCGGTTGCCCTTGGTAGCTAAGCGTGA 348
Db 208 CTGGCATTACAGAGCTTTCAGTTTTCAGATGGGCTTAAGCTGGTTAAGCTTAGAGGA 267
Qy 349 GCCTATATGGAAGAGCGGCTCTGCTGACTCTGGCAAGATGCT 392
Db 268 GAAGCTATGAGGATGCTGCTGATGCTGCAAAAAGTCAATGCT 311

```


